Received: 3 December 2020 | Revised: 30 September 2021 | Accepted: 2 October 2021

DOI: 10.1002/qj.4183

Quarterly Journal of the Royal Meteorological Society RMetS

INCOMPASS SPECIAL COLLECTION

The role of mid-tropospheric moistening and land-surface wetting in the progression of the 2016 Indian monsoon

Arathy Menon[1,2] | **Andrew G. Turner**[1,2] | **Ambrogio Volonté**[1,2] | **Christopher M. Taylor**[3,4] | **Stuart Webster**[5] | **Gill Martin**[5]

[1]Department of Meteorology, University of Reading, Reading, UK

[2]National Centre for Atmospheric Science, University of Reading, Reading, UK

[3]UK Centre for Ecology and Hydrology, Wallingford, UK

[4]National Centre for Earth Observations, Wallingford, UK

[5]Met Office Hadley Centre, Exeter, UK

Correspondence
A. Menon, Department of Meteorology, University of Reading, Reading RG6 6BB, UK
Email: arathy.menon@reading.ac.uk

Funding information
INCOMPASS project Weather, Grant/Award Numbers: NE/L01386X/1, NE/P003117/1, NE/L013819/1

Abstract

Accurately predicting the Indian monsoon is limited by inadequate understanding of the underlying processes, which feed into systematic model biases. Here we aim to understand the dynamic and thermodynamic features associated with the progression of the monsoon, using 2016 as a representative year, with the help of convection-permitting simulations of the Met Office Unified Model. Simulations are carried out in a 4 km resolution limited-area model, nested within a coarser global model. Two major processes thought to influence the northwestward progression of the monsoon are: (a) the interaction between the low-level monsoon flow and a mid-tropospheric dry-air intrusion from the northwest, and (b) land–atmosphere interactions. We find that the 4 km limited-area model simulates the mid-tropospheric moistening that erodes the northwesterly dry intrusion, pushing the northern limit of moist convection northwestwards. The surface soil moisture also plays a major role at the leading edge of the monsoon progression. The heavy rains associated with the local onset wet the soil, reducing the sensitivity of surface fluxes to soil moisture and weakening the land influence on further progression of monsoon rains. The 4 km model is tested with an alternative land-surface configuration to explore its sensitivity to land-surface processes. We find that the choice of soil and vegetation ancillaries affects the time-scales of soil moisture–precipitation feedback and the timing of diurnal convection, thereby affecting the local onset. We further compare these simulations with a parametrized convection run at 17 km resolution to isolate the effects of convective parametrization and resolution. The model with explicit convection better simulates the dynamic and thermodynamic features associated with the progression of the monsoon.

KEYWORDS

dry-air intrusions, Indian monsoon, land–atmosphere interaction, monsoon onset, monsoon progression, soil moisture

1 | INTRODUCTION

The Indian monsoon provides water for agriculture, industry and the livelihoods of more than a billion people in India. Changes in crop yield are directly influenced by the onset timing and the availability of the monsoon rains from June to September (Kumar *et al.*, 2004; Auffhammer *et al.*, 2012). The rain-fed agricultural sector plays a major role in the Indian economy. Hence, the forecasting of the Indian monsoon onset and progression is vital as it helps in water resource management, agricultural planning and thereby the gross domestic product of the country.

However, most general circulation models fail to accurately simulate the spatial and temporal variability of monsoon precipitation as significant systematic biases exist in these models. Atmosphere-only general circulation models (AGCMs), forced by observed sea surface temperatures (SSTs), generally simulate excess precipitation over the equatorial Indian Ocean and orographic slopes, such as the southern slopes of the Himalayas, and reduced precipitation over the Indian subcontinent (Sperber and Palmer, 1996; Gadgil and Sajani, 1998). Several studies have emphasised the importance of atmosphere–ocean coupled GCMs (AOGCMs) in simulating monsoon precipitation (Wu *et al.*, 2006); however, AOGCMs exhibit biases similar to those in AGCMs (Annamalai *et al.*, 2007; Sperber *et al.*, 2013). This indicates that the underlying causes lie within the atmospheric component of the models, while cold SST biases that arise from coupling act to aggravate these biases, thus delaying the onset (Levine and Turner, 2012; Levine *et al.*, 2013).

While global models perform reasonably at global scales, monsoon deficiencies may arise because the models fail to resolve important local- to-regional-scale processes (Karmacharya *et al.*, 2015). Studies using a variety of global and regional parametrized-convection models (or alternatively, regional models nested within global models) have tested the effect of increasing horizontal resolution. These studies have demonstrated improvements in the spatio-temporal distribution of Indian monsoon rainfall, especially over steep orography such as the Himalayas and Western Ghats (Ji and Vernekar, 1997; Kumar *et al.*, 2006; Johnson *et al.*, 2016). However, Johnson *et al.* (2016) showed that, while increasing the horizontal resolution globally from 200 to 40 km in the Met Office Unified Model (MetUM) has small, beneficial impacts, it does not solve the monsoon biases that are pervasive in GCMs.

Other findings suggest that many of the local errors are driven by large-scale remote biases that affect the monsoon mean state. Levine and Martin (2018) showed that a low resolution (50 km) RCM, forced at the lateral boundaries by reanalyses, could simulate monsoon lows and depressions better than a GCM at the same resolution, since the remote systematic biases are reduced in the RCM due to the smaller domain size, indicating that the mean state, more than (or as much as) the resolution, is important for simulating these phenomena when operating at the parametrized-convection scale.

A key factor limiting the reliability of the GCMs is the inability to accurately represent subgrid-scale processes. Convection, boundary-layer and land-surface processes are parametrized in the GCMs, yet many aspects of these processes are poorly represented. For example, parametrized convection results in many days of low rainfall accumulation and insufficient frequency of heavy rainfall (Yang and Slingo, 2001; Randall *et al.*, 2003; Holloway *et al.*, 2012; Marsham *et al.*, 2013a). Such models also show biases in the diurnal cycle of deep convection, with the convective maximum occurring at midday, rather than late afternoon, as observed (Yang and Slingo, 2001). A case-study in West Africa using MetUM has shown that simulations using convective parametrization results in the largest errors in components of the water cycle, associated with the diurnal cycle and the location of convection (Birch *et al.*, 2014). Although the MetUM with parametrized convection reproduces the key rain-forming sea-breeze circulation over the Maritime Continent region, the model fails to respond realistically to the sea-breeze circulation (Birch *et al.*, 2015).

Parametrized convection also introduces some biases in the response of the atmosphere to land-surface conditions. Global atmospheric models with convective parametrization fail to capture the observed preference for afternoon rain over drier soils (Taylor *et al.*, 2012). Soil moisture heterogeneity triggers mesoscale circulations, resulting in convergence and triggering of convection, an important mechanism that has been observed over India (Barton *et al.*, 2019). At daily time-scales, parametrized simulations show enhanced rainfall over wetter soils, resulting in an erroneous tendency for a positive feedback in those models (Taylor *et al.*, 2013), which indicates that typical parametrization schemes might be overly sensitive to evapotranspiration.

Convection-permitting models do not rely on convective parametrization schemes and typically allow for a more accurate representation of the surface and orography due to their increased resolution (Prein *et al.*, 2015). Willetts *et al.* (2017) showed that, despite the grid spacing exceeding the scale of individual convective clouds, convection-permitting simulations lead to improvements in the general rainfall biases over India, with more intense rainfall and a later peak in the diurnal cycle of convective precipitation over land. The convection-permitting model shows greater skill than the global model over East Africa, especially on sub-daily time-scales and for storms over land (Woodhams *et al.*, 2018). Over West Africa, it is clear

that thermodynamic biases from parametrized convection translate onto the continental-scale monsoon dynamics (Marsham *et al.*, 2013a). Simulations at 4 km resolution using explicit convection better capture the rainfall rate than the 12 km simulations using parametrized convection, the biases arising mainly from the parametrization rather than the resolution itself (Holloway *et al.*, 2012). Convection-permitting models also capture the negative feedback between soil moisture and convective triggering, indicating the ability of these models to simulate the mesoscale circulations that generate convergence (Taylor *et al.*, 2013). Hence, high-resolution simulations with explicit convection are potentially useful for a better understanding of the Indian monsoon and the involvement of land–atmosphere interactions in its development and progression.

Modelling land–atmosphere fluxes in a forecast or climate simulation over a large region is inherently challenging due to the heterogeneity of the landscape. Surface fluxes are sensitive to the nature of the vegetation, properties of the soil, as well as soil moisture, which itself evolves in response to antecedent rainfall, evaporation and drainage. Land-surface parametrizations are required to simulate these processes based on often grossly simplified maps of soil and vegetation. As a result, climate models struggle to capture realistic evaporation dynamics (Gallego-Elvira *et al.*, 2019), particularly in regions like India, where irrigation is widespread and profoundly affects surface fluxes (Bhat *et al.*, 2019).

One of the major features of the Indian monsoon is the northwestward advance of the rains in a direction perpendicular to the direction of the mean monsoon flow. There are several theories suggested for this northwestward propagation. Ananthakrishnan *et al.* (1968) suggested that the slow seasonal intensification of the monsoon trough and the associated increase in pressure gradient and frictional convergence results in this northwestward progression. Wang (2005) suggested that the meridional progression of the monsoon is associated with the northward progression of intraseasonal oscillations. The eastward-propagating boreal summer Madden–Julian Oscillation (MJO) also has a substantial influence on the Indian monsoon onset and progression (Wheeler and Hendon, 2004; Wu *et al.*, 2006).

Using observational and reanalysis data, Parker *et al.* (2016) suggested that the northwestward progression of the monsoon is modulated by mid-tropospheric dry-air intrusions from the northwest. As the monsoon progresses, these dry northwesterlies are moistened from the southeast by shallow cumulus and congestus clouds which form when convection tries to penetrate the freezing level (Johnson *et al.*, 1996). This increases the relative humidity (RH) of the ambient air near the freezing level and results in more cloud formation and deepening of this high-RH layer, eventually eroding the dry-air intrusions from beneath and progressively from the southeast. Later, Menon *et al.* (2018) analysed the dry-air intrusion mechanism in the initialised coupled seasonal hindcasts of GloSea5-GC2. Analysis of cloud fraction and specific humidity data from the model supports the mid-tropospheric moistening mechanism suggested by Parker *et al.* (2016). Some studies analyse the relative roles of large-scale dynamics (such as intraseasonal oscillations as mentioned in the above paragraph) and atmospheric moistening by congestus clouds in triggering convection on diurnal scales (e.g., Hohenegger and Stevens, 2013). However, in this study we will not discuss the relative roles of these mechanisms in the onset progression.

Parker *et al.* (2016) explained that the dry-air intrusion is removed by the arrival of a moister and warmer airmass. Volonté *et al.* (2020) elaborated on this finding by showing that the progression of the monsoon towards northwest India is a non-steady process which is modulated by the balance of the interaction between two airmasses with markedly different moist static energy (MSE), that is, the northwesterly intrusion of dry air from western and central Asia and the moist low-level southwesterly flow from the Arabian Sea. The local onset of the monsoon is associated with the progression of the sharp boundary between these two airmasses, and thus by the passage from low-θ_e to high-θ_e conditions in the local atmosphere. AGCM experiments in Chakraborty *et al.* (2006) have shown that the large-scale monsoon onset over India occurs only after the surface MSE reaches a threshold value and the large-scale vertical ascent reaches the mid-troposphere. This study suggests that the instability of the atmosphere, and the organised large-scale ascent are two necessary conditions that lead to the Indian summer monsoon onset. Their experiments removing the west Himalayan orography show that the resulting cold-air intrusion makes the atmosphere more stable, thereby delaying the onset.

Another major feature that could affect the northwestward monsoon progression is the land-surface wetting by pre-monsoon showers. High-resolution model simulations using the Advanced Research Weather Research and Forecasting Model (ARW-WRF) show that an increase in soil moisture to the immediate north of the Indian monsoon onset front, from lighter rains emanating in the anvils of deep convective clouds, results in an increase in buoyancy and development of new clouds towards the north of the front which is steered to the north and eventually to the northwest by a divergent circulation (Krishnamurti *et al.*, 2012). Hence, this progression is highly sensitive to the parametrization of soil moisture and non-convective anvil rains. Land–atmosphere models have suggested that northern India features strong land–atmosphere coupling on seasonal scales (Koster *et al.*, 2004).

As discussed above, it is well established that convective parametrization has many drawbacks such as producing rainfall with a diurnal peak too early in the day, which could exacerbate model biases at different time-scales. Also, GCMs typically do not include a representation of irrigation, which is widespread across the Indo-Gangetic Plains. It is still unclear how the Indian monsoon onset mechanism behaves at convection-permitting resolutions and how land-surface inputs will alter the onset processes in the model. The Interaction of Convective Organisation with Monsoon Precipitation, Atmosphere, Surface and Sea (INCOMPASS) project aims to deliver a significant improvement in the capability of Indian monsoon forecasting by gathering observational data and by improving the understanding of physics and dynamics associated with the Indian monsoon. To this end, we undertook a field campaign in India (Turner *et al.*, 2020) during summer 2016, followed by modelling of the 2016 Indian monsoon using the Met Office Unified Model (MetUM; Cullen, 1993; Brown *et al.*, 2012). The seamless modelling approach of the MetUM enables the implementation of improvements in physical processes, once identified, to a broad range of spatial and temporal scales (Martin *et al.*, 2010). We also conducted model experiments with different land-surface configurations, to be discussed in this paper. This study focuses on the local processes that affect Indian monsoon progression using 2016 as an example, given the observational and modelling efforts already performed for the year.

Hence, in this study, we compare 2016 monsoon onset in the MetUM with explicit and parametrized convection to understand the impact of convective parametrization, if any. We also compare the onsets under alternative land-surface specifications to understand the effect of different land-surface types and associated fluxes in the progression of the monsoon. We aim to understand the impact of resolution and parametrization on the role of mid-level dry-air intrusions and land-surface wetting on the northwestward progression of the monsoon using high-resolution, season-long model simulations for 2016. Section 2 provides a description of the model, the observational and reanalysis data used and the methods. Section 3 discusses the effect of mid-tropospheric moistening and land-surface wetting in the progression of the 2016 monsoon. Section 4 focuses on the role of the land surface in the diurnal cycle of precipitation. Results are summarised in Section 5.

2 | DATA AND METHODS

This section describes the model experimental set-up and the other data and methods used in this study.

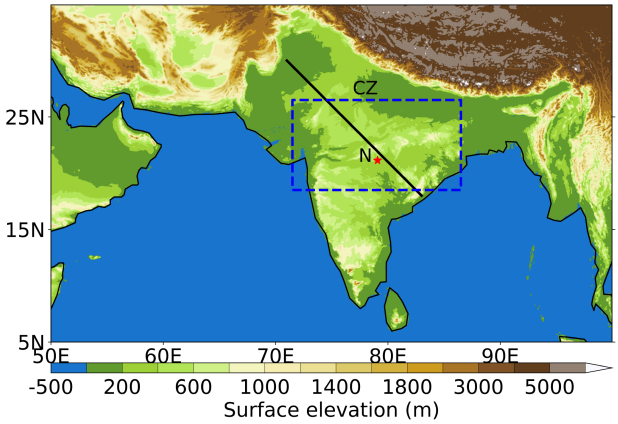

FIGURE 1 The LAM domain, showing orography (colour shading) from the 4 km model. The back line indicates the cross-section along which various thermodynamic parameters are calculated in Figures 6–8. The blue box shows the monsoon core zone. Nagpur is marked by the red star [Colour figure can be viewed at wileyonlinelibrary.com]

2.1 | Model description and experimental set-up

All the simulations in this study are based on the MetUM, a non-hydrostatic fully compressible model with atmospheric dynamics solved by a semi-implicit, semi-Lagrangian dynamical core (Davies *et al.*, 2005). Limited-area model (LAM) simulations were performed for the summer months of 2016 from 01 May to 30 September 2016 at 4 and 17 km horizontal resolutions for the Indian region 50°–100° E and 5° N–35° N (the nested model domain is shown in Figure 1). The 4 km LAM has a vertical resolution of 80 levels with a 38.5 km lid, while the 17 km LAM has a vertical resolution of 70 levels with an 80 km lid. The model uses a rotated latitude–longitude horizontal grid with Arakawa C-grid staggering and a hybrid-height vertical coordinate with Charney–Philips staggering. The LAM is nested within the global model at N768 resolution ($\Delta x \approx 17$ km). The driving global model, which is a prototype version of the Global Atmosphere 7.0 (GA7) and Global Land 7.0 (GL7) configurations (Williams *et al.*, 2018), is initialized daily at 0000 UTC from the global analysis. The LAM is allowed to run freely within the driving global model. The lateral boundary conditions for the LAM is provided by the global model, while SSTs are updated daily using the Operational Sea Surface Temperature and Sea Ice analysis (OSTIA; Donlon *et al.*, 2012).

The dynamics of both the driving global model and the LAM are based on Even Newer Dynamics for General Atmosphere Modeling of the Environment (END Game; Wood *et al.*, 2014). The driving global model and the 17 km LAM use a prognostic cloud fraction

TABLE 1 Summary of the model simulations performed, including the labels identifying the experiments in the study

Model	Convective parametrization	Large-scale cloud scheme	Vegetation cover data
17 km LAM (LAM17-I)	Yes	PC2	IGBP
4 km LAM (LAM4-I)	No	Smith (1990)	IGBP
4 km LAM (LAM4-C)	No	Smith (1990)	CCI

and condensation (PC2; Wilson *et al.*, 2008) scheme whereas the 4 km model uses the diagnostic Smith (1990) scheme. Boundary-layer turbulence is parametrized using a blended scheme (Boutle *et al.*, 2014) which combines the one-dimensional (1D) scheme (Lock *et al.*, 2000) with 3D Smagorinsky mixing. Stratton *et al.* (2018) provide more details of the 4 km model configuration. The major differences between the LAM experiments performed are summarised in Table 1.

The Joint UK Land Environment Simulator (JULES) model (Best *et al.*, 2011; Clark *et al.*, 2011) is used to calculate the fluxes of heat, moisture and momentum into the atmosphere. Subgrid-scale heterogeneity is introduced through a tile approach using nine surface tiles that includes five plant functional types (PFTs): broad leaf trees, needle leaf trees, C3 grass, C4 grass, shrubs and four non-vegetated surface types: urban, inland water, bare soil and land ice. Surface fluxes are calculated separately on each tile. Soil processes are calculated using a four-layer scheme and these four soil layers have a thickness (from top) of 0.1, 0.25, 0.65 and 2 m. The hydraulic conductivity of unsaturated soil is calculated from soil moisture using the Van Genuchten (1980) relationship.

The land-surface boundary conditions are provided through the land ancillaries. Two experiments were performed with the convection-permitting 4 km model by using two different land ancillaries. One simulation used the land–sea mask, soil properties and vegetation parameters from the International Geosphere and Biosphere Programme (IGBP) dataset (Loveland *et al.*, 2000) which has a resolution of about 1 km globally. The second simulation used an alternative land-surface configuration from the European Space Agency Climate Change Initiative (CCI) ancillaries (Poulter *et al.*, 2015), which has an original resolution of 300 m. CCI uses up-to-date spatial and thematic land cover information that is consistently mapped over time. In addition, we imposed spatially uniform soils across the domain. This idealised approach avoids the development of spurious gradients in surface fluxes associated with highly uncertain soil properties and their impact on evapotranspiration in the model. The surface forcing for the 17 km simulation was provided by the IGBP ancillaries. Table 2 shows parameters which are different between the two ancillaries.

TABLE 2 List of parameters which are different in CCI and IGBP land ancillaries

Land-surface parameters
Volume fraction of condensed water in soil at wilting point of plants
Volume fraction of condensed water in soil at critical point of plants
Volume fraction of condensed water in soil at saturation
Clapp–Hornberger "B" coefficient (soil drainage types)
Soil thermal conductivity
Soil hydraulic conductivity at saturation
Soil thermal capacity
Saturated soil water suction
Snow-free albedo of soil
Soil carbon content
Fractions of surface types
Leaf Area Index of plant function types (monthly varying climatology)
Canopy height of plant function types (monthly varying climatology)

Over India (Figure 2), CCI shows a reduction in C3 grasslands compared to IGBP, however C4 grasslands are more dominant in the CCI ancillary. CCI has more trees and less bare soil over northwest India near Rajasthan than the IGBP ancillary (Figure 2). Regions with more C3 and C4 grass in comparison to trees and bare soil can dry down faster after a rainfall event, since evapotranspiration decay time-scales are shorter for grasslands (Martínez-de la Torre *et al.*, 2019).

2.2 | Observational data and methods

In order to compare the model rainfall with observations, we use a satellite–gauge merged daily precipitation dataset (Mitra *et al.*, 2009) which merges the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) satellite estimates with the India Meteorology Department (IMD) gridded rain gauge

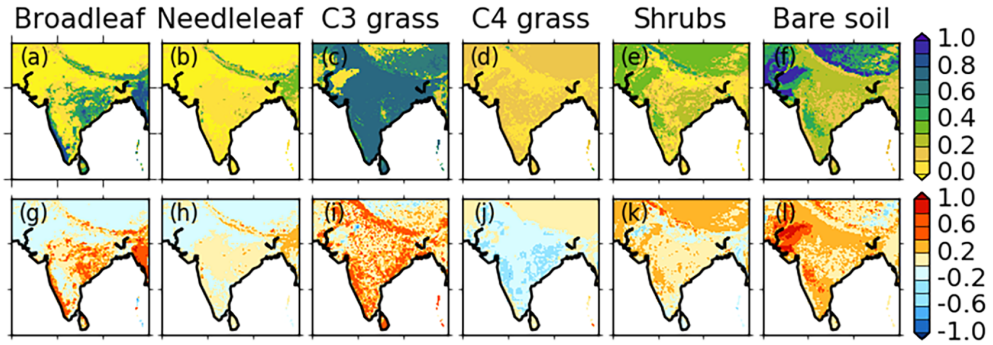

FIGURE 2 (a–f) Land cover fractions from IGBP ancillaries for the plant functional types broad leaf trees, needle leaf trees, C3 grass, C4 grass, shrubs and surface bare soil, respectively. (g–l) show the differences between the CCI and IGBP land fractions (IGBP minus CCI) [Colour figure can be viewed at wileyonlinelibrary.com]

dataset. This merged product has a horizontal resolution of 0.5° × 0.5°. The diurnal cycle of precipitation in models is compared with the 3-hourly satellite-based rainfall estimates from the Climate Prediction Center morphing method (CMORPH; Joyce *et al.*, 2004). Relative humidity and winds from the ERA-Interim re-analysis dataset (Dee *et al.*, 2011) for 2016 at different pressure levels are also used to compare with the model output. ERA-Interim has a horizontal resolution of 0.75° × 0.75°.

To understand the effect of the land surface on moist convection, we use evaporative fraction (EF), the ratio of latent heat to the sum of latent and sensible heat. Evaporative fraction is often nearly constant during daytime (e.g., Gentine *et al.*, 2011). It is therefore a useful diagnostic of the partition between sensible and latent heat fluxes.

3 | PROGRESSION OF THE 2016 INDIAN MONSOON

We first analyse the role of model resolution, parametrization and land-surface boundary conditions in the northwestward progression of the monsoon.

3.1 | Model performance at simulating progression of the rains

The progression of the 2016 monsoon from the merged satellite–gauge data is shown in Figure 3a1–i1. Each panel shows pentad-average rainfall from 01 June to 15 July. During the beginning of the season, the rainfall was mainly distributed over southwestern India and over the Bay of Bengal where rainfall was greater than 18 mm·day^{-1}. As the season continued, the rainfall progressed northwestward, covering the whole country by around mid-July. The climatological monsoon onset date over Kerala (southwest India) is 01 June, but in 2016, the observed onset was on 08 June. Even though the onset was delayed by a week, with a rapid progression during the last week of June the monsoon covered the whole of the country by around 13 July (Prasad *et al.*, 2018).

Significant MJO activity was observed in the Indian Ocean from 12 to 18 June 2016 (Shao *et al.*, 2018); the MJO is known to perturb the Indian monsoon onset (Wheeler and Hendon, 2004). Suppressed convection over India during 11–15 June was associated with this MJO phase which enhances convection over the equatorial Indian Ocean and suppresses convection over India (Figure 3c1).

As the MJO started moving eastwards, the Bay of Bengal became more convectively active (Figure 3d1). After the MJO moved to the Maritime Continent by around 19 June, a northwest–southeast tilted rain band appeared over India and the surrounding seas (Figure 3e1). This is the northward-propagating Oceanic Tropical Convergence Zone (OTCZ) as suggested by Pai *et al.* (2011), which propagated further northwards in the following weeks. This northwest–southeast band is also a signal of the Boreal Summer Intraseasonal Oscillation (BSISO) which was active during this period (Fletcher *et al.*, 2020). This enhanced the northward progression of the monsoon and contributed to its rapid progression during the last week of June, irrespective of the delayed onset. A monsoon depression which was present over the northern Arabian Sea from 27 to 29 June (IMD, 2016) also contributed to the heavy rainfall over this region in the 26–30 June pentad (Figure 3f1).

Another major synoptic feature that affected the progression of the 2016 monsoon was a low pressure system that formed over the Bay of Bengal during the end of June, which later moved over land and became a depression on 06 July (Turner *et al.*, 2020). During the beginning of July (Figure 3g1), rainfall was restricted mainly to the eastern parts of India and the northern Bay of Bengal by the low pressure system. Heavy precipitation over central India during 6–10 July was associated with the monsoon depression. The monsoon progressed further northwestwards after the movement of the depression westwards (Figure 3i1).

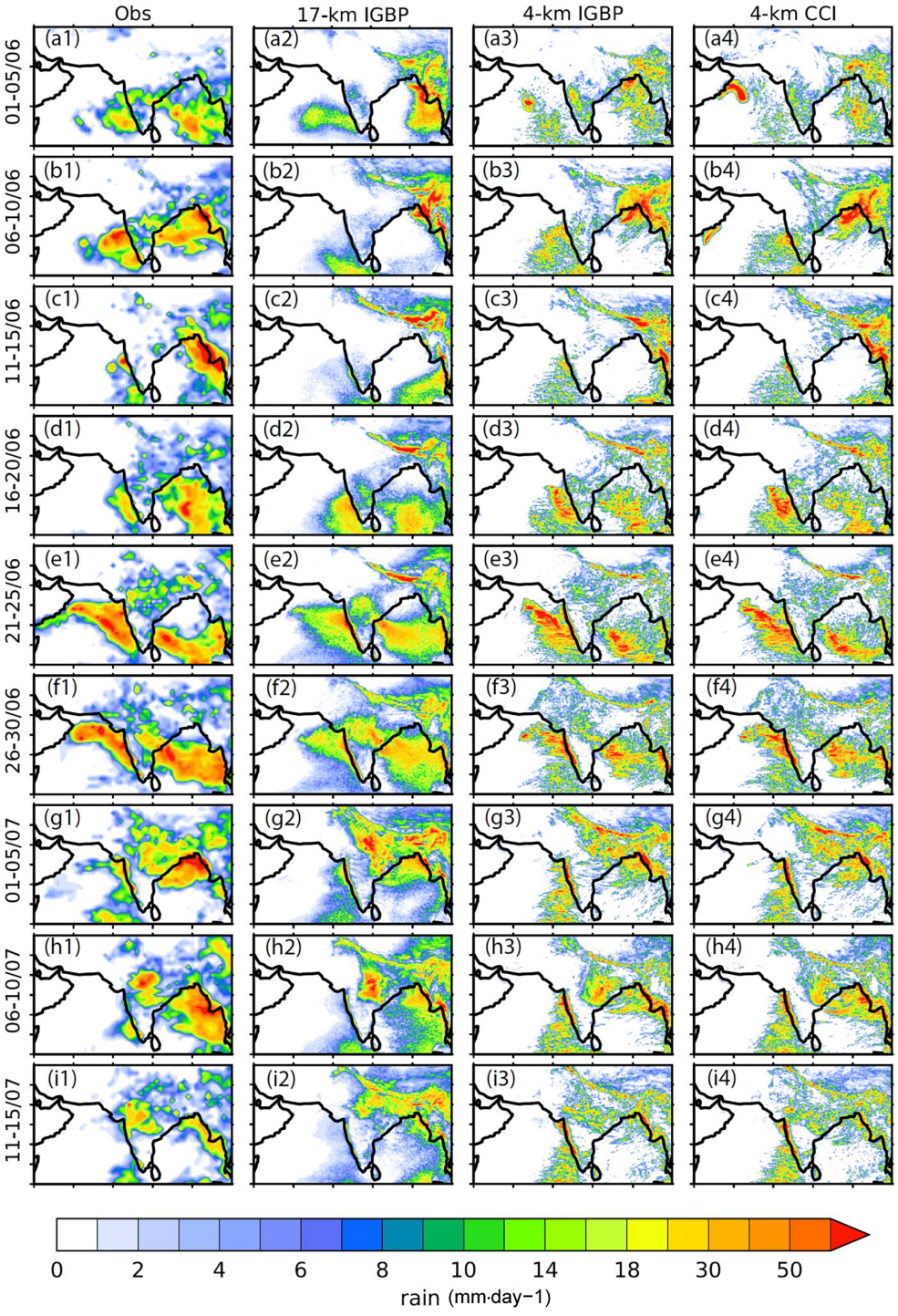

FIGURE 3
Pentad-average rainfall from the beginning of June to mid-July 2016 from the (a1–i1) IMD–GPM merged precipitation dataset, (a2–i2) 17 km simulation using IGBP land ancillary, (a3–i3) 4 km simulation using IGBP ancillary and (a4–i4) 4 km simulation using CCI ancillary. LAM4-I and LAM4-C data have been re-gridded to 17 km resolution [Colour figure can be viewed at wileyonlinelibrary.com]

The 17 km simulation (which uses the IGBP land ancillaries) significantly underestimates the total rainfall rate along the MJO-associated rain bands and in the northern Bay of Bengal (Figure 3a2–i2).

The monsoon depression is shifted slightly westwards in the 17 km simulation (Figure 3h2) compared to the 4 km simulations, making the location of the depression closer to observations in the 17 km simulations. The intensity of the rainfall associated with the monsoon depression is higher in the 17 km model than in the 4 km models. The peak in daily rainfall during the beginning of July in Figure 4 over the core monsoon zone was associated with the monsoon depression. Both the 17 km model and the IMD merged precipitation capture rainfall peaks of about 23 mm·day^{-1} associated with the depression, whereas the 4 km models capture only half of this precipitation (about 10 mm·day^{-1}). Much of the rainfall associated with the depression is stratiform due to the presence

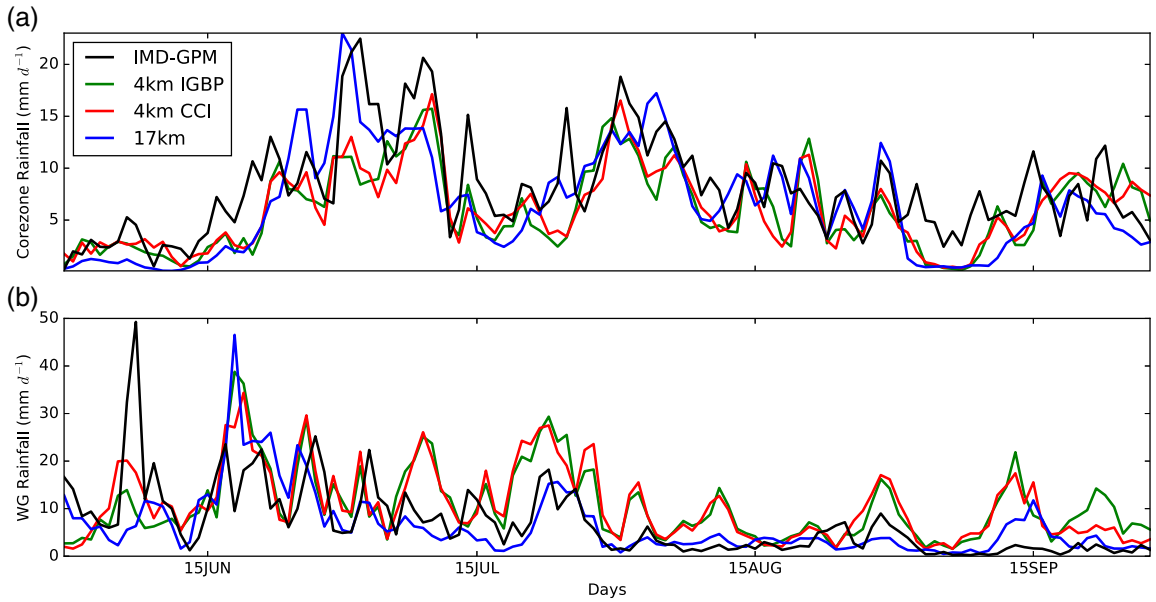

FIGURE 4 Daily precipitation (mm·day^{-1}) during 1 June to 30 September 2016 from the IMD–GPM merged dataset and 17 km, 4 km IGBP and 4 km CCI simulations averaged over: (a) monsoon core zone (71.5°–86.5°E,18.5°–26.5°N) and (b) Western Ghats (73.5°–77.5°E, 7.5°–15°N) regions [Colour figure can be viewed at wileyonlinelibrary.com]

of large-scale low-level convergence (Hunt and Turner, 2017). The model with parametrized convection captures this rainfall well compared to observations.

The 4 km models with both land ancillaries are able to capture the northwestward progression of the monsoon (Figure 3) and the rainfall associated with the two major features (MJO and depression). The pentad-mean rainfall patterns look quite similar in the two experiments (Figure 3). Figure 4 shows that the evolution of daily mean rainfall during the monsoon season over the core monsoon zone and the Western Ghats are very similar in the two 4 km simulations with different land ancillaries, however their diurnal cycles are systematically different as we shall see later in Section 4. The position of the depression is shifted slightly eastwards in both simulations (Figure 3h3, h4) compared to the observations. Both simulations have a negative rainfall bias over the central Bay of Bengal and a positive bias over northeastern India during the beginning of the season compared to observations (not shown).

It is clear that the models simulate the observed northwestward progression of the 2016 monsoon quite well. The 4 km simulations capture the general spatial pattern of the rainfall better when compared to observations. The location of the monsoon depression and associated rainfall are captured well by the model with parametrized convection, which is remarkable considering large size of the LAM domain and that the LAM is free-running within the global model. We now examine two of the major factors (mid-tropospheric moistening and land-surface wetting) which could be affecting the northwestward progression of the monsoon, using 2016 as an example.

3.2 | Mid-tropospheric moistening

As discussed in the Introduction, climatologically the progression of the Indian monsoon is affected by a mid-tropospheric dry-air intrusion from the northwest (Parker *et al.*, 2016). Figure 5 shows the spatial pattern of dry-air intrusions using the vertical distribution of relative humidity and winds for 2016 in the reanalysis and models. At 850 hPa, towards the end of May, winds were westerly over southern India, which are the mean monsoon winds. These winds carry moisture to southern India from the Arabian Sea, resulting in relative humidity above 60%. However, the simulations are drier over the eastern parts of India and northern Bay of Bengal than in the reanalysis. At 600 hPa, there were strong northwesterly winds of the order of 5–15 m·s^{-1} that carried dry air (RH < 50%) from the Afghanistan region, all the way to the Bay of Bengal (Figure 5a). This represents the mid-tropospheric dry-air intrusion in 2016, as observed by the research flights of INCOMPASS (Turner *et al.*, 2020), and is captured by all the simulations. However, in LAM17-I, these northwesterlies become zonally oriented over eastern India at 600 hPa, and the dry-air layer over northern Bay of Bengal at 850 hPa is drier in the 17 km simulation than in the two 4 km models.

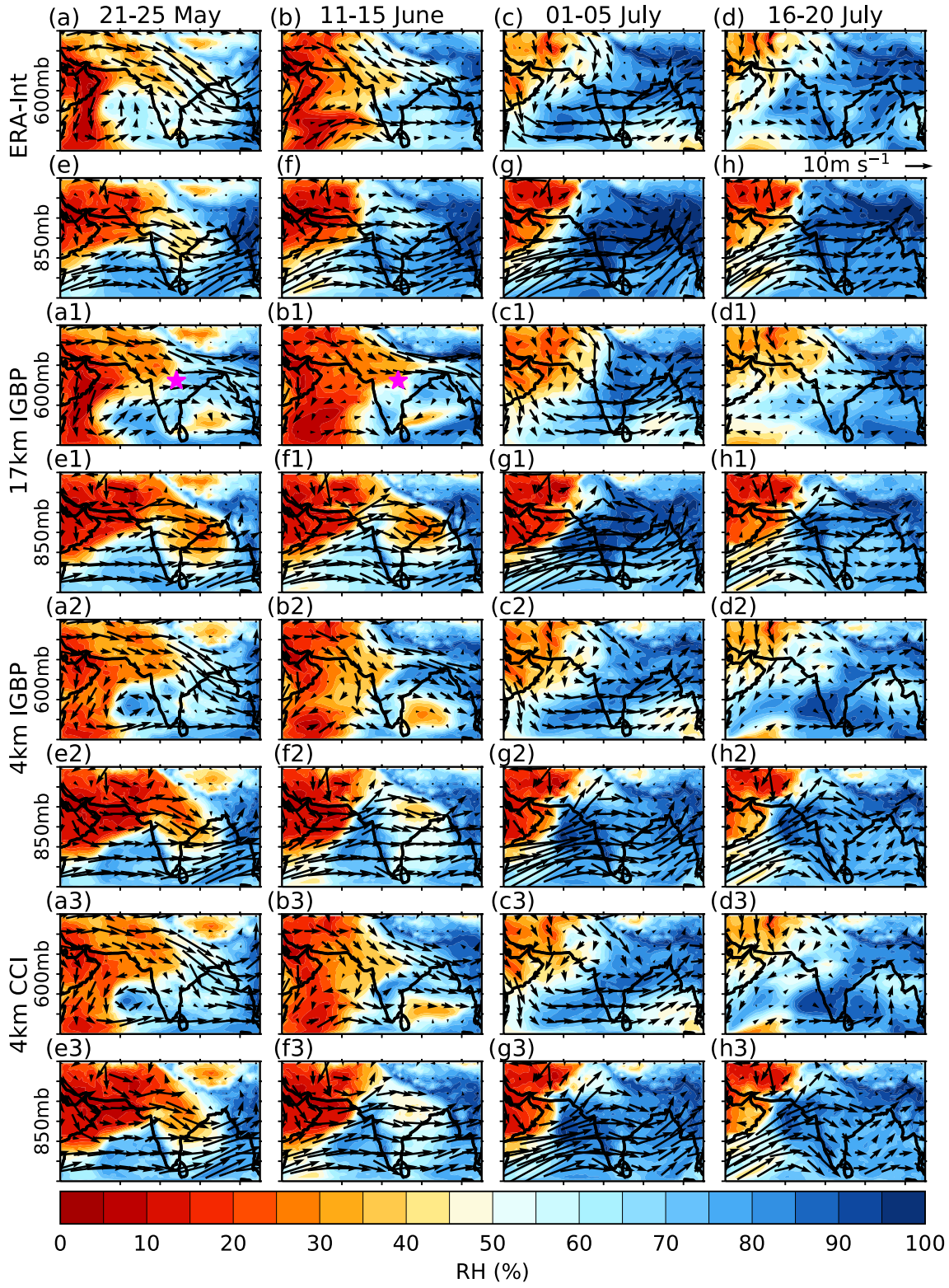

FIGURE 5 Mean relative humidity and winds at 600 and 850 hPa averaged over 5-day periods from (a–h) ERA-Interim reanalysis, (a1–h1) the 17 km model, (a2–h2) the 4 km model with IGBP ancillaries and (a3–h3) the 4 km model with CCI ancillaries. All model data are re-gridded to the ERA-Interim grid for comparison. Magenta star in (a1) and (b1) shows the location of Nagpur, which is shown here to aid the explanation of Figure 9 in Section 3.2 [Colour figure can be viewed at wileyonlinelibrary.com]

In June, the intensity of the dry-air intrusion at 850 hPa is overestimated (drier) and is more widespread in the 17 km model than in ERA-Interim and the 4 km models. Towards the end of May, a well-defined cyclonic vortex is evident at 600 hPa over the eastern Arabian Sea in both 4 km simulations. However, the vortex is located over the western Arabian Sea and is more meridionally aligned in both LAM17-I and the reanalysis. Steering of the winds by this mid-level vortex and the dry-air intrusion results in the formation of a north–south ridge along the western coast of India at 850 hPa.

As time progressed, the mid-level dry-air intrusion receded and the westerlies strengthened in depth from the lower to mid-troposphere, establishing a strong monsoon flow. By around mid-July (Figure 5 last column), the westerlies over southern India started to weaken and the northwesterlies at upper levels became northeasterlies and flowed towards the northwestern Arabian Sea. By mid-July, the monsoon covered the whole of the country with relative humidity greater than 60% over most of the land.

We now look at the detailed thermodynamics of the 2016 monsoon progression and dry-air intrusion in comparison to climatology. Figure 6 shows the vertical distribution of several thermodynamic parameters around 20 May, 15 June and 10 July, along a northwest–southeast cross-section (as shown in Figure 1) from the 4 km simulation that uses IGBP ancillaries. The layers of higher water vapour mixing ratio (r) near the surface in Figure 6a–c represent the monsoon layer. As time progressed, the monsoon layer extended towards the northwest and deepened from the southeast. This increase in moisture content can be explained by the strengthening and deepening of the monsoon westerly winds and the shallowing of the mid-tropospheric northwesterlies as the monsoon progressed. The evaporative fraction (purple line, Figure 6) increased towards the northwest as the season progressed. The very high values of EF on 9–11 July (Figure 6c) appear unrealistic in comparison to observations from a range of sites across India (Bhat *et al.*, 2019).

We now analyse the vertical distribution of RH (Figure 6d–f), following Parker *et al.* (2016), as it plays a significant role in triggering cumulus convection. Around 20 May, a high relative humidity wedge with RH above 50% developed in the southeast between 500 and 600 hPa, which is the freezing level over India at this time. This high-RH wedge is associated with detrainment from the cumulus clouds as the convection tries to penetrate the freezing level, as suggested by Johnson *et al.* (1996) and Parker *et al.* (2016).

The humid layer was deeper and sloped downwards towards the northwest in June (Figure 6e), showing a marked increase in relative humidity at lower levels (e.g., 800 hPa) associated with winds carrying moisture from the Arabian Sea towards northwest India. At higher levels, the dry-air intrusions from the northwest resulted in lower values of RH (< 30%) in the northwest above 700 hPa. By 10 July, as the monsoon continued its progression, the entire atmospheric column below 400 hPa over southeastern and central India showed RH above 60%.

We further examine the vertical structure of the equivalent potential temperature (θ_e) to understand the evolution of the dry-air intrusion and mid-tropospheric moistening. The layer of lowest θ_e near 500 hPa around 20 May (Figure 6g) shows the dry-air intrusion. The dry layer is deeper towards the northwest. As the monsoon progressed, this dry layer was moistened from the southeast by cumulus and cumulus congestus clouds, as is evident from the increase in mid-level moisture in Figure 6e. However, this initial moistening at mid-levels (Figure 6d) does not increase the θ_e in the mid-troposphere (Figure 6g). As discussed in some earlier studies (e.g., Madden and Robitaille, 1970), θ_e can be used as an analogue for static stability. Hence, we can say that the initial moistening at mid-levels (500–600 hPa) from shallow convection (Figure 6d) does not change the MSE at these levels. However later, as the moist monsoon airmass with high MSE reached the dry layer, the θ_e at these levels increased from ∼ 340 to ∼ 350 K (Figure 6h, i). As a result, the dry layer became shallower and was eroded from the southeast by 10 July (Figure 6i).

Saturation equivalent potential temperature (θ_{es}) indicates the convective regime. A large vertical gradient in θ_{es} above 850 hPa (Figure 6j) around 20 May was associated with the low dry-static stability of the dry-air intrusion from the northwest. As the monsoon progressed, the θ_{es} near the surface decreased towards the southeast. This is associated with cooling of the surface and the boundary layer because of increased soil moisture from monsoon rains and the depression-associated rainfall. Higher soil moisture allowed more energy to go into evaporation, resulting in a cooling of the ground and the overlying air. Figure 7 shows the difference between LAM4-I and LAM4-C. The difference between water vapour mixing ratios (r) from both IGBP and CCI simulations is less than 1 g·kg^{-1} at upper levels, but in July in the boundary layer, r is about 4 g·kg^{-1} more in CCI than in IGBP at grid points along the centre of the cross-section. This is consistent with locally enhanced evaporative fraction in CCI compared to IGBP in the aftermath of rain from the monsoon depression. Boundary-layer humidity anomalies of similar amplitude and spatial scale were observed in response to soil moisture features during the INCOMPASS field campaign (Barton *et al.*, 2019).

Mid-tropospheric moistening is weaker in LAM4-C than in LAM4-I in May and June (Figure 7d, e; the figure shows the values of IGBP minus CCI). This is

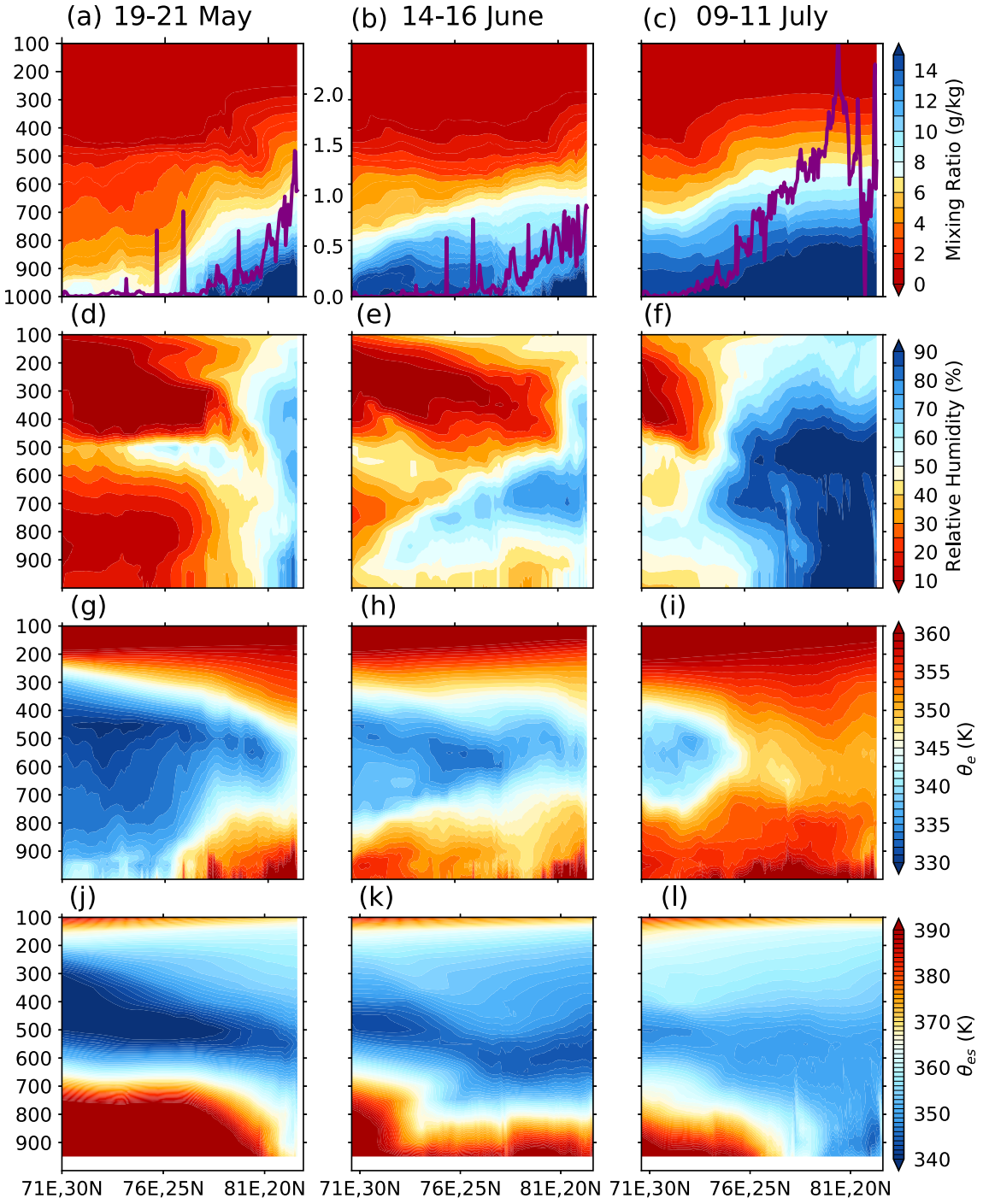

FIGURE 6 Northwest to southeast (left to right) vertical sections of (a–c) water vapour mixing ratio (g·kg^{-1}), (d–f) relative humidity (%), (g–i) θ_e (K) and (j–l) θ_{es} (K) averaged over 3 days around (a, d, g, j) 20 May, (b, e, h, k) 15 June and (c, f, i, l) 10 July from the 4 km simulation that uses IGBP land ancillaries. The axis of the cross-section is given in Figure 1. The purple lines in (a–c) show the evaporative fraction along the cross-section, as measured by the right *y*-axis in (a). [Colour figure can be viewed at wileyonlinelibrary.com]

also evident from lower values of θ_e in LAM4-C, in the mid-troposphere, along the regions of mid-tropospheric moistening (Figure 7g, h). Other major differences in both simulations are in the boundary layer, due to different land ancillaries in these simulations. θ_{es} near the surface is lower in LAM4-C towards the southeast, showing an increase in boundary-layer moisture in this region from higher evaporative fraction. More details of the evaporative fraction and land–atmosphere interactions in these simulations are discussed in the next section.

Figure 8 shows the vertical distribution of these thermodynamic parameters in LAM17-I, which uses

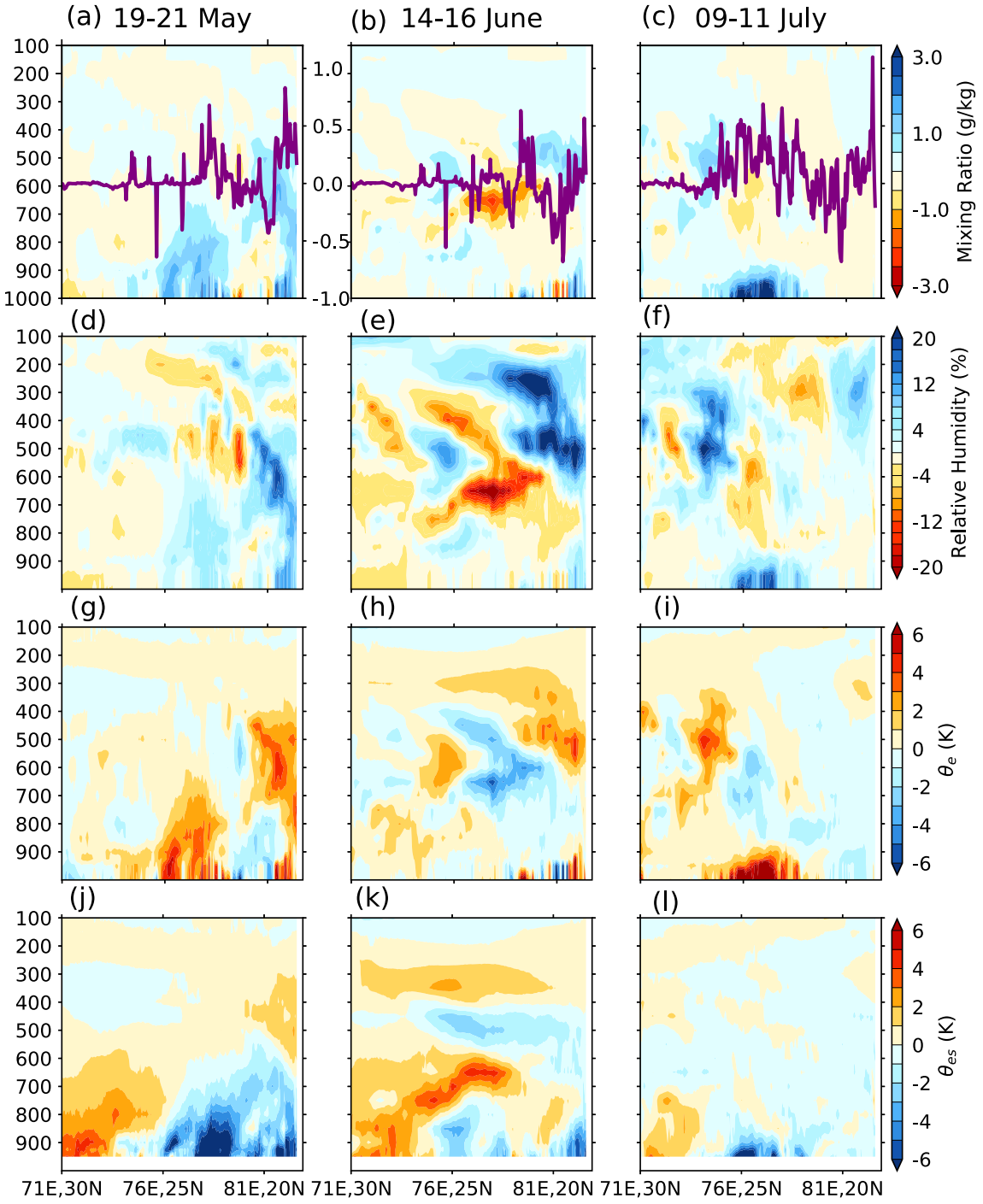

FIGURE 7 As Figure 6, but showing differences between the 4 km simulations using CCI and IGBP land ancillaries (IGBP minus CCI). The purple lines in (a–c) show the difference in evaporative fraction between the two experiments, as measured by the right *y*-axis in (a) [Colour figure can be viewed at wileyonlinelibrary.com]

IGBP land ancillaries. The 17 km model captures the mid-tropospheric cloud formation around 20 May and its extension towards the northwest with the progression of the monsoon. However, the 17 km model is not as good as the 4 km models in capturing the gradual erosion of the dry-air intrusions by mid-tropospheric moistening from the southeast. Mid and low levels are drier on 15 June than on 20 May (shown by lower mixing ratios) and this is reflected in the lower values of θ_e around 15 June (Figure 8h). This is also reflected in the larger differences between θ_e and θ_{es} at low levels and is associated with the higher θ_{es} on 15 June. The lower θ_e layer that depicts the dry-air intrusion is much drier in the 4 km simulation than in the 17 km simulation in May. Generally, the magnitude

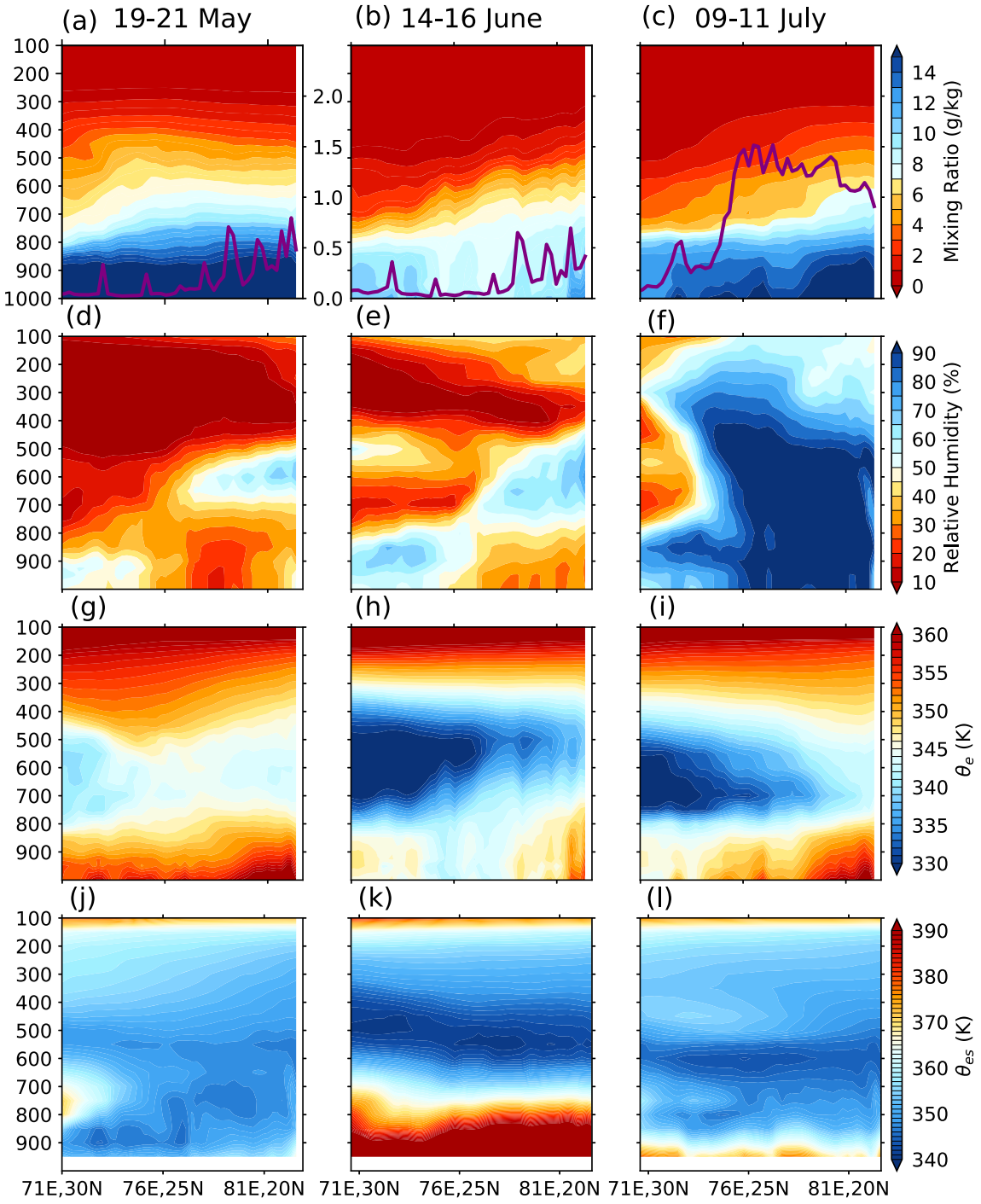

FIGURE 8 As Figure 6, but showing results from the 17 km IGBP simulation. [Colour figure can be viewed at wileyonlinelibrary.com]

of the differences in these thermodynamic parameters are higher between the parametrized and explicit simulations at different resolutions, compared to the two explicit 4 km simulations with different land ancillaries (Figure 7). Even though the two IGBP simulations use the same land ancillary, the evaporative fraction is lower in the 17 km simulation towards the southeast than in the 4 km simulation, reflecting the differences in rainfall between the two simulations.

For a better understanding of the thermodynamics of the interaction of dry-air intrusions with monsoon progression, we look further into the time evolution of some of these thermodynamic parameters at Nagpur (Figure 9), a location in central India (red star in Figure 1), following Parker *et al.* (2016). The local onset of the monsoon at Nagpur in 2016 was on 15 June. Until about 20 days prior to the local onset, the atmospheric column in the 4 km simulation using the CCI ancillary was very

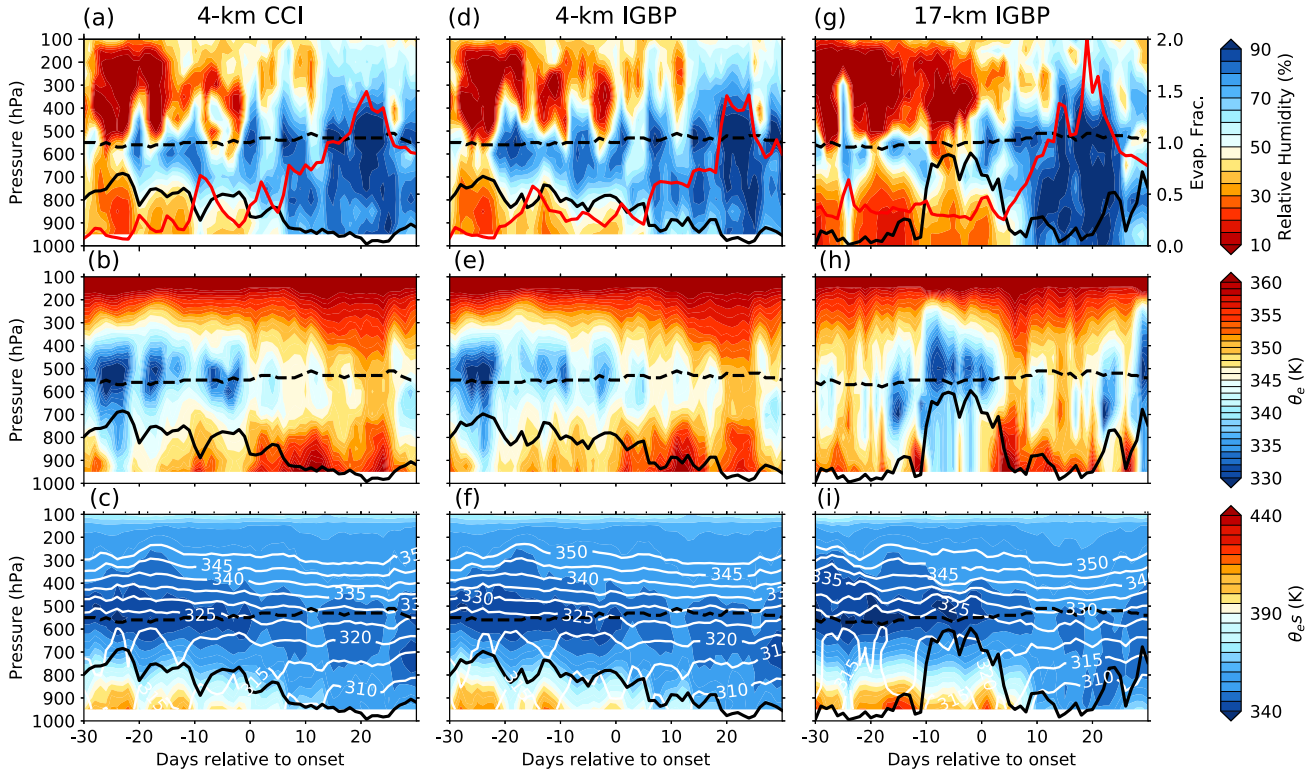

FIGURE 9 Time–pressure section at the grid point closest to Nagpur of (a, d, g) relative humidity (%), (b, e, h) θ_e (K), (c, f, i) θ_{es} (K) with virtual potential temperature θ_v overplotted (white contours, K). (a–c) are from the 4 km simulation with CCI ancillaries, (d–f) from the 4 km simulations with IGBP ancillaries, and (g–i) from the 17 km simulation. In all panels, the black dashed line represents the freezing level ($T = 0\,°C$) and the black solid line shows the Lifting Condensation Level (LCL). In (a, d, g) the red lines represent the evaporative fraction, with axis at right of (g). The profiles are computed for days relative to the local circulation onset at Nagpur on 15 June. [Colour figure can be viewed at wileyonlinelibrary.com]

dry over Nagpur with RH<40%, which suppressed the development of deep cumulus clouds. At the end of this period, a layer of high RH formed near the freezing level about 20–22 days prior to the monsoon onset at Nagpur (Figure 9). About 10 days prior to the onset, the evaporative fraction increased and the humid layer extended towards the surface. This humid layer gradually deepened and about 15 days after the local onset the entire tropospheric column became humid with relative humidity exceeding 60%. The LCL lowered in altitude as the monsoon progressed in time. The simulation using the IGBP ancillary exhibits a similar relationship between EF and LCL, but with differences in timing of high EF/low LCL periods (Figure 9 second column). However, the 17 km simulation (Figure 9 third column) shows a deep humid layer around the freezing level that extends toward the surface (accompanied by an increase in evaporative fraction) about 25 days prior to the onset. This humid layer later becomes shallower for a few days and then deepens again after the onset. This is also evident from the maps of relative humidity; during 21–25 May, the relative humidity over Nagpur (denoted by purple star) at 600 hPa is higher than in the 11–15 June period (Figure 5a1, b1). The LCL is lower until about 10 days prior to onset; later the LCL rises to about 600 hPa and again lowers following the onset. The freezing level is between 500 and 600 hPa in all the simulations. The evaporative fraction increases following the onset up to about 20 days and decreases afterwards. This post-onset increase in evaporative fraction is similar in the two 4 km simulations, but the increase is higher in the 17 km simulation.

The lowest values of θ_e near 600 hPa represent the dry-air intrusion as mentioned above. Around 10 days prior to the onset, the EF increased (Figure 9a) and around that time θ_e also increased near the surface (Figure 9b), consistent with an increase in low-level MSE. This increase in MSE at low levels could decrease the stability of the atmosphere, making it conducive for convection, as suggested by Chakraborty and Agrawal (2017). At the mid-troposphere, the RH started to increase about 20 days prior to the onset from detrainment and shallow convection at the freezing level (Figure 9a). However, this moistening was not associated with an increase in θ_e. But as the local onset approached, the high-MSE monsoon flow acted to increase θ_e through the depth of the entire troposphere, including the mid-levels. The low-θ_e layer then became

shallower as the mid-tropospheric RH increased, eroding the dry-air intrusion as the onset approached (Figure 9b,e). These features are consistent with the theory of Parker *et al.* (2016) and the results shown by Menon *et al.* (2018): as the onset approaches, shallow clouds started to form at the altocumulus layer near the freezing level and these clouds penetrated to higher levels, moistening the tropospheric profile over time, thereby eroding the dry-air intrusions. The simulation using IGBP ancillaries also shows a similar vertical structure. However, in the 17 km simulation, the dry layer of lowest θ_e values appears at Nagpur only about 15 days prior to the onset. The deep humid layer that is seen about 25 days prior to the onset in the 17 km simulation as mentioned above, might be moistening the dry-air intrusions at mid-levels prior to 15 days. From Figure 5 it is clear that in LAM17-I, the mid-levels at Nagpur are drier during the days closer to the onset than in May. In other words, in the 17 km simulation, the erosion of dry-air intrusions by mid-tropospheric moistening is more unsteady than in the two 4 km models. This sudden moistening and drying in the 17 km model might be due to the tendency of the convective parametrization scheme to produce intermittent deep convection with little or no shallower convection in between (Martin *et al.*, 2017). Stirling and Stratton (2012) also found such intermittency in convection in the diurnal cycle over land.

A month prior to the local onset, θ_{es} values were higher near the surface, but about 10 days prior to the onset, θ_{es} below 700 hPa started decreasing (Figure 9c,f). This cooling of low-level air preceding the onset could result in increasing dry static stability around the onset. The near-surface θ_{es} reduced by about 20 °C following the onset. Despite the cooling of the low-level air, its θ_e increases since the low-level moisture increases. Widely spaced contours of virtual potential temperature (θ_v) near the surface show the presence of a deep adiabatic layer. Ten days prior to the onset, the θ_v contours started to become closer together, showing the transition into a pseudoadiabatic state, as suggested by Parker *et al.* (2016). IGBP and CCI 4 km simulations give similar results. The 17 km simulation does not capture the near-surface decrease in θ_{es} prior to the onset.

To summarise, both the mid-tropospheric dry-air intrusion and the moist southwesterly monsoon flow with markedly larger MSE are important for the progression of the monsoon, since the onset is characterised by the θ_e "jump" that signals the changes in the airmass (as outlined by Volonté *et al.*, 2020). The 4 km models are able to capture the thermodynamics associated with this mechanism, as shown by the time–pressure section at Nagpur (Figure 9a–f).

3.3 | Land-surface wetting by pre-monsoon showers

We now look into the role of the land surface as the soil moisture immediately to the north of the onset front could affect the progression of the monsoon (Krishnamurti *et al.*, 2012), with an emphasis on the sensitivity of this mechanism to model resolution, convective parametrization and land ancillaries.

A Hovmöller diagram along a 5-degree meridional cross-section through central India shows the northward progression of the monsoon rains (Figure 10 contours). The surface-layer (0–10 cm) soil moisture clearly responds strongly to precipitation in both 4 km simulations. However, in the CCI run, post-rain surface soil moisture is lower than for IGBP. This is primarily because of differences in soil properties between the two simulations, which affect both infiltration and evaporation. The CCI run is prescribed with uniform sandy soils with a saturated hydraulic conductivity an order of magnitude higher than typical values in the IGBP simulation. This results in more rapid movement of soil water down the profile after rain. In addition, the CCI soil also has a lower "critical point" than typical IGBP soils, and this parameter defines the soil moisture stress functions used in the model to constrain evapotranspiration (Best *et al.*, 2011). As a result, surface fluxes in the CCI simulation are more sensitive to rainfall when the soils are relatively dry. These two effects combine to produce the weaker temporal fluctuations in surface soil moisture evident for CCI in Figure 10, but at the same time, stronger water fluxes occur both further down into the soil, and back into the atmosphere.

Figure 11 shows the extent to which recent rainfall patterns control the partition between sensible and latent heat fluxes across the monsoon core zone. During May the soil is largely dry, and the energy balance is dominated by sensible heat flux (Figure 11b). Any localised showers during this period, for example on 22 May, create patches of wet soil which can evaporate strongly for several days. In this highly water-limited regime, there is a strong control on EF from recent rain and surface fluxes, as quantified by the spatial correlation between 3-day antecedent rainfall accumulations and EF (Figure 11a). Note that this correlation weakens with rainfall accumulation period (not shown) as the surface dries out. Throughout June, rainfall gradually wets the soil profile, and area-mean sensible heat flux falls approximately four-fold (Figure 11b). This period coincides with a pronounced seasonal maximum in spatial variability in sensible heat flux due to the patchiness of rain falling on relatively dry soils. However, by July even locations which have experienced little rain in previous days can maintain modest rates of evapotranspiration, and this

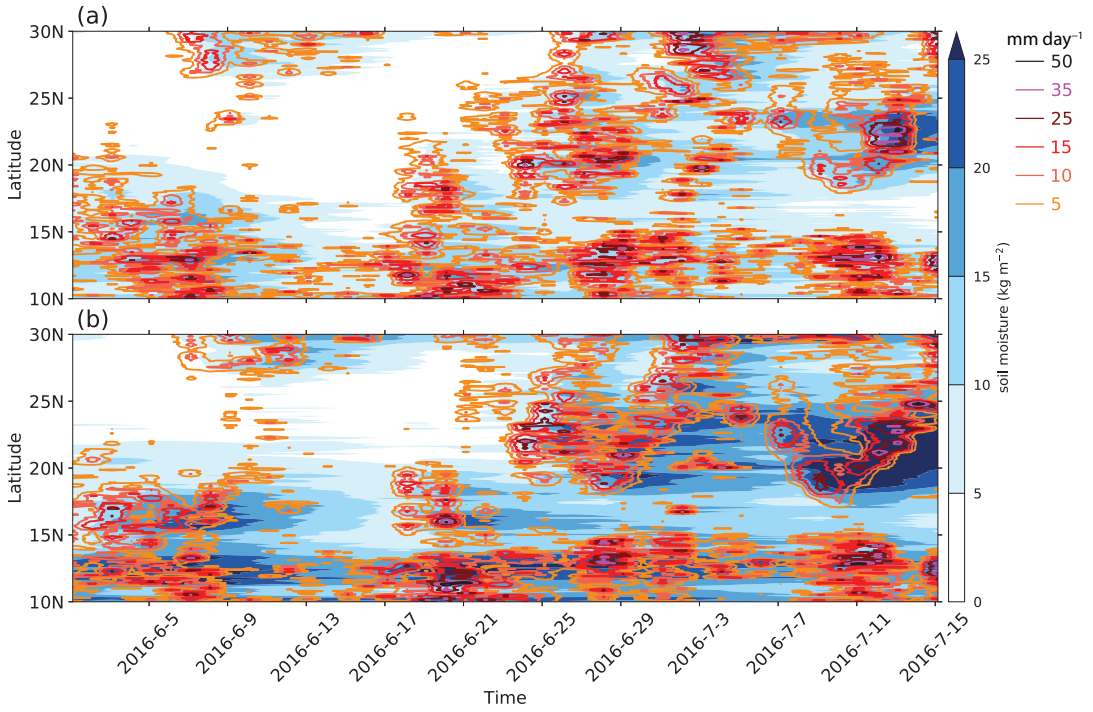

FIGURE 10 Latitudinal variation of rainfall (mm·day^{-1}; contours) and top 10 cm soil moisture (kg·m^{-2}; shading with interval 5 kg·m^{-2}) averaged over the longitudes 75°–80° E for the 4 km simulation using (a) CCI and (b) IGBP ancillaries. The darker (lighter) contours show higher (lower) rainfall [Colour figure can be viewed at wileyonlinelibrary.com]

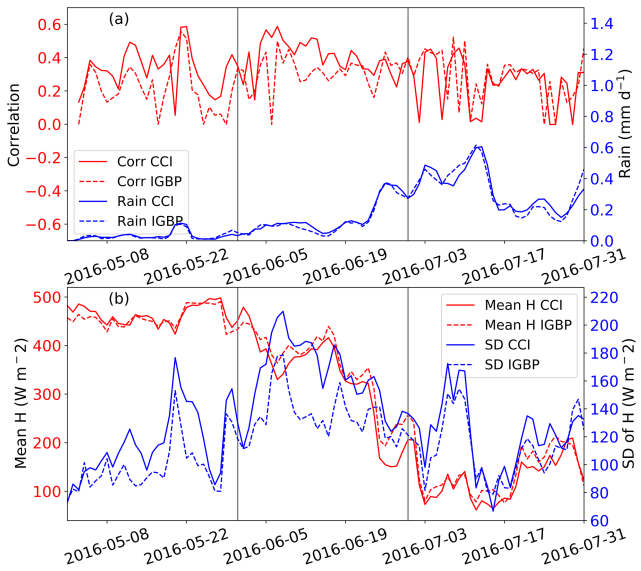

FIGURE 11 Time evolution of (a) 3-day mean rainfall (blue lines), and the spatial correlation between 3-day antecedent rains and 0630 UTC (1200 IST) evaporative fraction (red lines); (b) area-mean sensible heat flux (H, red lines) and spatial standard deviation of the sensible heat flux (blue lines) from the two 4 km simulations with IGBP and CCI land ancillaries at 0630 UTC (1200 IST) over the monsoon core zone. The grey vertical lines separate the months May, June and July [Colour figure can be viewed at wileyonlinelibrary.com]

widespread wetting effectively suppresses strong spatial variations in sensible heat. The transition from high (May) to low (July) mean sensible heat flux via a regime with strong spatial variability in surface heating (June) is evident in both simulations. However, in the CCI simulation spatial variability in sensible heat almost always exceeds that in the IGBP run, particularly in the June wetting-up phase (Figure 11b).

In summary, CCI has stronger flux sensitivity to antecedent rain when soil is very water stressed (May, June) and this drives stronger spatial variability in the sensible heat flux H. In July, CCI still has stronger spatial variability of H than IGBP for much of July which, accompanied by a much more unstable atmosphere, could produce more daytime rain. We next try to understand the effect of these land ancillaries on the diurnal evolution of moist convection in the two 4 km simulations.

4 | EFFECT OF THE LAND SURFACE ON THE CONVECTIVE DIURNAL CYCLE

4.1 | Diurnal cycle of precipitation

Before looking at the detailed processes on the impact of the land surface, we summarise the diurnal cycle of rainfall

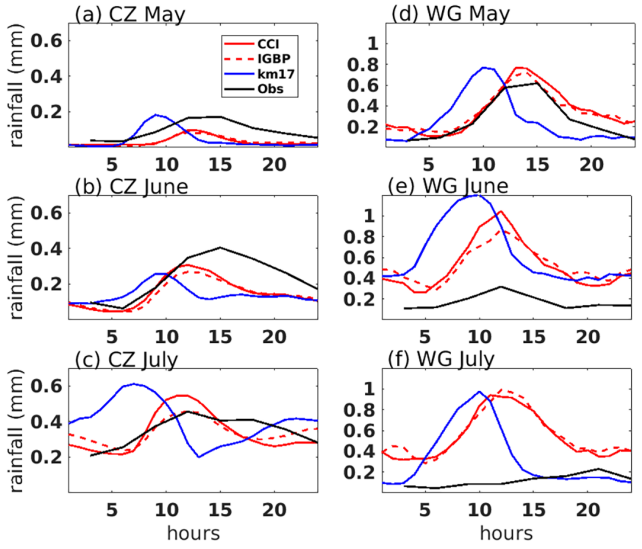

FIGURE 12 Diurnal cycle in rain (mm·hr^{-1}) in CMORPH analyses, 17 km, 4 km IGBP and 4 km CCI simulations for the months (a, d) May, (b, e) June and (c, f) July averaged over (a–c) the monsoon core zone (71.5°–86.5°E,18.5°–26.5°N) and (d–f) the Western Ghats region (75°–77°E, 8.5°–14°N). Only land points are considered and all the times are in UTC [Colour figure can be viewed at wileyonlinelibrary.com]

for each of the model configurations for May, June and July (Figure 12). This is performed over the core monsoon zone which provides the majority of rain over central India and in the Western Ghats region of steep orography that intercepts the monsoon flow as it passes from the Arabian Sea. During May, the diurnal maximum rainfall for the monsoon core zone in the parametrized simulation is around 0.2 mm·hr^{-1}, consistent with the observed value, whereas in the explicit simulations it is about 0.1 mm·hr^{-1}. During June, the explicit simulation with CCI ancillaries simulates a slightly higher diurnal peak than the two IGBP simulations. The peak precipitation amount in the models in June is about 0.1 mm·hr^{-1} less than the observed values. The diurnal cycle of precipitation is largest in July, compared to the other two months, for the monsoon core zone.

Over the Western Ghats, which is a region of high orography, the diurnal amplitude is larger than in the core zone. The values of May and July rainfall are comparable in the parametrized and explicit simulations (0.8 mm). However in June, the diurnal maximum in the parametrized simulation is much higher than in the explicit simulations.

In both regions, the model with parametrized convection (blue lines in Figure 12) peaks too early in time compared to the explicit simulations. This is a common issue with convective parametrization, as discussed in the Introduction. The parametrized run has a diurnal maximum at around 0600 UTC during July for the core zone, which is about 6 hr too early relative to observations. For other months and over the Western Ghats, the parametrized convection captures a diurnal maximum in rainfall at around 0900 UTC. The timing of the diurnal maximum in the explicit runs is between 1000 and 1500 UTC, which is similar to observations.

The diurnal peak in precipitation occurs a few hours later over the Western Ghats than in the core monsoon zone in all simulations and the observations (except in July). This could be because, during the day, the land is heated faster than the Arabian Sea to its west, which could trigger onshore winds and increased convective available potential energy (CAPE), resulting in convective storms over the mountains to the east in the late evenings. In the core monsoon region, heating of the land during the day might be helping the air parcels to overcome the convective inhibition; with the help of increased CAPE, this could result in the convective peak occurring earlier in the day.

4.2 | Spatial controls on convective initiation

As discussed above, a major feature of the diurnal peaks of precipitation over the monsoon core zone (Figure 12a–c) is that the two explicit simulations have a similar diurnal peak in May. Later, in June and July, the peaks diverge in amplitude with LAM4-C simulating a slightly higher peak in June and a much higher peak in July than LAM4-I. Over the Western Ghats, the rainfall diurnal peaks in LAM4-I and LAM4-C converge from June to July (Figure 12d–f).

Why do strong differences in the diurnal cycle develop between the two 4 km runs? To answer this, we look at where daytime precipitation is triggered in these simulations relative to the spatial distribution of surface fluxes. We compute the spatial correlation between daytime precipitation and antecedent EF over sub-areas of the core zone (by excluding the coastal and high orography areas to remove their effects on mesoscale precipitation). This sub-region provides us with 40 1° × 1° grid boxes to calculate the daily spatial correlations. For each day, we correlated the 0600 UTC evaporative fraction in each grid box to succeeding rains only when (a) there was no rainfall on that day prior to 0600 UTC (1130 IST) anywhere in that region, and (b) it rained at some point from 0700 to 1400 UTC (1230–1930 IST) in at least 25% of the grid points in a grid box in that region. This process produces a distribution of up to 40 correlation coefficients per day which, in the absence of a consistent feedback from EF to rainfall, would be symmetric about zero. Figure 13 shows the distribution of correlation coefficients for 19 May to

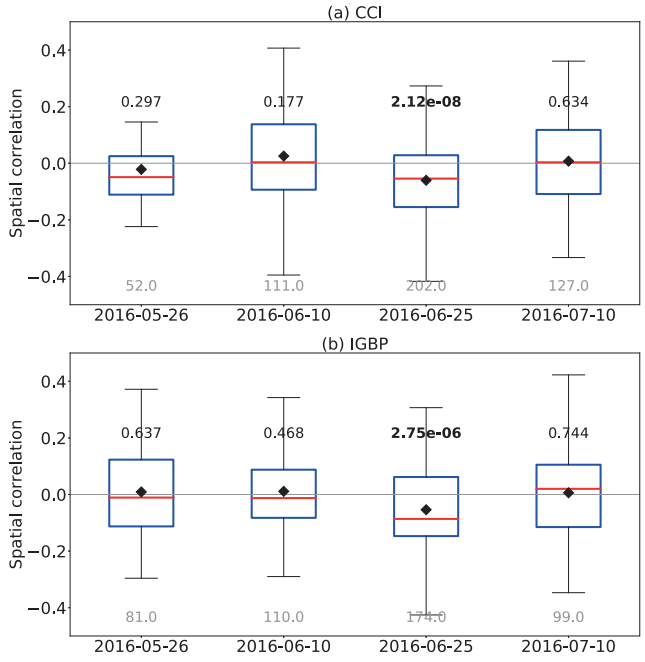

FIGURE 13 Daily spatial correlations between 0600 UTC evaporative fraction and 0700–1400 UTC rainfall for the period 19 May to 18 July 2016. Each box represents the daily spatial correlations aggregated over 15 days for the core zone sub-regions. The whiskers represent 25th and 75th percentiles. The red lines represent the median and the black diamonds show means for each 15-day period. Grey numbers at the bottom of each box represent the number of 1° boxes sampled in 15-day periods which satisfy the conditions mentioned in Section 4. A two-tailed *t*-test is performed to understand the significance of the mean of these correlations. *p*-values are given at the top of each box (black numbers) and those which are significant at or below the 5% level (0.05) are shown in bold. (a) is for LAM4-CCI and (b) is for LAM4-IGBP [Colour figure can be viewed at wileyonlinelibrary.com]

18 July aggregated over 15-day blocks (to increase the data points in the distribution for more robust statistics). As we go from the 15-day period centred on 10 June to the 15-day period centred on 25 June, there is a significant deviation in the mean correlation coefficient from zero to negative values (significant *p*-values from a two-tailed *t*-test are shown as bold numbers in Figure 13). Later, during the next 15 days centred on 10 July, the spatial correlation between daytime precipitation and antecedent EF goes back to zero. This shows that, as the onset approaches the core zone, the diurnal rains are more frequently triggered over the drier regions by increased sensible heat flux, which results in this negative correlation; once the onset is established and the soil gets wetter, the feedback weakens, which is shown by the zero correlation between precipitation and antecedent EF. This is consistent with the observed soil moisture–precipitation feedback found by Taylor *et al.* (2012). Spatial variations in soil moisture and sensible heat trigger boundary-layer circulations and cloud development in response to the diurnal cycle in surface heating, as shown by Barton *et al.* (2019) for India and Rieck *et al.* (2014) for the midlatitudes.

The mean correlations are significant at 95% confidence levels for the 15-day period from 18 June to 2 July, which covers the period of monsoon onset over the core zone. During 2016, the monsoon onset over the core zone was between 15 and 26 June according to IMD (2016). Also, from Figure 11 it is clear that the rainfall over the core zone started to increase from 18 June. This means that the effect of soil moisture heterogeneity on diurnal rainfall was more prominent during the arrival of the onset. The soil properties of CCI result in stronger spatial variability in sensible heating than IGBP, which will favour earlier triggering of convection in CCI, consistent with its enhanced daytime rainfall (Figure 12). The CCI simulation continued to produce more spatial variability in sensible heat flux than IGBP for much of July, and it is these same periods that are responsible for the enhanced rainfall diurnal cycle evident in Figure 12. It is only during the monsoon depression and its aftermath that both the mean and spatial variability in sensible heat flux becomes weak in both models, and this coincides with damped daytime precipitation (not shown).

5 | SUMMARY

Models used for monsoon forecasting exhibit systematic biases. Previous studies have shown that many of these biases develop quickly, and are not improved substantially by increased model resolution. This suggests that they may be related to poor parametrization schemes. Parametrizations of convection, boundary-layer and land-surface processes are highly sensitive to large-scale drivers and underlying surface conditions which may not be well-known or represented. Such parametrizations, in their current formulations, may also be unable to represent detailed processes such as the diurnal cycle of convection and the effect of the land surface on convective triggering. Studying the processes in detail at convection-permitting scales, despite not properly resolving convective processes, may improve understanding and ultimately be beneficial for the improvement of models and thereby for monsoon forecasting. In this study, we have documented the nested suite experiments performed using the MetUM under the INCOMPASS project to understand the physics and dynamics behind the northwestward progression of the monsoon. Understanding the advance of the monsoon is necessary for understanding and correcting some of the known biases in weather and climate models. Nested suite simulations were carried out at 4 and 17 km horizontal

resolutions and the 4 km simulation was also tested with two different land ancillaries.

The interaction between mid-level dry-air intrusion from the northwest and monsoon flow with high MSE is an important control on the local onset and advance of the monsoon. Parker *et al.* (2016) and Menon *et al.* (2018) showed that, as the onset progresses northwestwards, the mid-tropospheric dry layer is moistened from below by detrainment from shallow cumulus and congestus clouds from the southeast. These clouds form when convection tries to penetrate the freezing level, as shown by Johnson *et al.* (1996). The dry layer becomes much shallower towards southeast India, making the profile closer to moist adiabatic, providing favourable conditions for deep cumulus convection. Increased moistening of the free troposphere thereby pushes the northern limit of moist convection northwestwards.

We have observed this dry-air intrusion in 2016 during the field campaign. The nested suite simulations in this study show that the 4 km models using explicit convection better capture the role of mid-tropospheric dry-air intrusions in the northwestward progression of the monsoon than the 17 km model with parametrized convection, since convection plays a major role in this process. The 4 km models with explicit convection appear to represent the dry-air intrusion and the gradual moistening of the troposphere as the monsoon onset progresses. The 17 km model with parametrized convection does not capture well this gradual erosion of the dry-air intrusion of 2016. It rather shows a nonlinear moistening of the mid-troposphere as the monsoon progresses.

The model with explicit convection also throws light onto the role of land-surface interactions in the progression of the monsoon. During the local onset of the monsoon over a particular region, the onset or pre-monsoon showers result in an increase in soil moisture heterogeneity which increases the spatial variability in evaporative fraction. Increased gradients in sensible heat generate daytime mesoscale circulations which favour earlier triggering of rain. However, when the soil becomes wetter in July, the evaporative fraction becomes less sensitive to soil moisture. The 4 km models capture this effect of pre-monsoon showers on soil moisture and convective initiation along the leading boundaries of the monsoon progression.

In the 4 km simulations, the strength of the soil moisture–precipitation feedback depends on the soil and vegetation ancillaries used. LAM4-I and LAM4-C feature different time-scales of evaporative response to rainfall. LAM4-C has a strong but short-lived evaporative response, whereas LAM4-I has a more muted and slower build-up of evaporation as the monsoon progresses. As rain is patchy and the surface is initially dry in the convergence zone, the soil moisture heterogeneity created by antecedent rains translates to stronger spatial variability in EF and H in LAM4-C compared to LAM4-I. This spatial variability in the surface fluxes feeds back on the timing of diurnal convection through generation of circulations, which bring forward the triggering of convection to earlier in the day in LAM4-C. Barton *et al.* (2019) finds enhanced rains over the drier surfaces, which is physically consistent with an earlier onset of convection over locally deeper boundary layers, probably driven by mesoscale circulations. Hence identifying and using better land-surface ancillaries is essential for the improvement of monsoon forecasts. However, the effect of model resolution and convective parametrization is dominant.

From the arguments made in this study, we propose the following summary regarding the two mechanisms in the progression of the monsoon:

1. The mid-tropospheric moistening plays a major role in the progression of the monsoon.
2. Land-surface wetting moistens the boundary layer and helps in the initial progression of the onset rains. However, once the rainfall wets the soil over appreciable depth, surface fluxes become less sensitive to the variations in soil moisture, and mid-tropospheric moistening plays the major role in further progression of the monsoon. This could be tested in further experiments, beyond the scope of this study, in which evapotranspiration responses to soil moisture variations are switched off (e.g., set to a constant), to test for any impact on monsoon progression.

In this study, we have not found any obvious evidence linking the land surface to moistening the upper levels of the free troposphere on a large scale, but it does not rule out that mechanism. Chakraborty and Agrawal (2017) showed that an increase in moisture near the surface increases low-level MSE, which decreases the stability of the atmosphere and makes it conducive for convection. We have clear evidence from this study and from Barton *et al.* (2019) that the land surface is moistening the boundary layer, which further moistens the free troposphere via detrainment from shallow convection.

Encouragingly, the initial results from this study suggest that the spatial pattern of the monsoon progression and the dynamics associated with the dry-air intrusions and land-surface processes are better represented in the 4 km model with explicit convection than in parametrized convection at 17 km. A proper understanding of the convective life cycle of monsoon precipitation and its evolution during the monsoon onset is a prerequisite for eventually improving parametrization schemes. High-resolution simulations for multiple years would also be beneficial

for checking the statistical significance of the mechanisms highlighted in this study.

AUTHOR CONTRIBUTIONS

Arathy Menon: formal analysis; investigation; methodology; visualization; writing – original draft. **Andrew G. Turner:** conceptualization; funding acquisition; investigation; methodology; supervision; writing – review and editing. **Ambrogio Volonté:** methodology; visualization; writing – review and editing. **Christopher M. Taylor:** methodology; visualization; writing – review and editing. **Stuart Webster:** resources; writing – review and editing. **Gill Martin:** methodology; visualization; writing – review and editing.

ACKNOWLEDGEMENTS

This work was funded by the INCOMPASS project (NERC grant numbers NE/L01386X/1, NE/P003117/1, NE/L013819/1), a joint initiative between the UK Natural Environment Research Council and the Indian Ministry of Earth Sciences. GM is supported by the Weather and Climate Science for Service Partnership (WCSSP) India, a collaborative initiative between the Met Office, supported by the UK Government's Newton Fund, and the Indian Ministry of Earth Sciences (MoES). We thank Sonja Folwell for providing us the CCI land ancillaries used in the simulations of this study. The ERA-Interim data used in this study can be downloaded from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim. The model data are availabale upon request and we will be making it publicly available in the near future via the CEDA/BADC archive. We thank the Editors and two anonymous reviewers for their comments on this work.

ORCID

Arathy Menon https://orcid.org/0000-0001-9347-0578
Andrew G. Turner https://orcid.org/0000-0002-0642-6876
Ambrogio Volonté https://orcid.org/0000-0003-0278-952X
Christopher M. Taylor https://orcid.org/0000-0002-0120-3198
Gill Martin https://orcid.org/0000-0003-0851-6020

REFERENCES

Ananthakrishnan, R., Srinivasan, V., Ramakrishnan, A.R. and Jambunathan, R. (1968). Synoptic features associated with the onset of southwest monsoon over Kerala, in *India Meteorological Department Forecasting Manual, Part III*. Poona: IMD.

Annamalai, H., Hamilton, K. and Sperber, K.R. (2007) The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations. *Journal of Climate*, 20, 1071–1092.

Auffhammer, M., Ramanathan, V. and Vincent, J.R. (2012) Climate change, the monsoon, and rice yield in India. *Climatic Change*, 111, 411–424.

Barton, E.J., Taylor, C.M., Parker, D.J., Turner, A.G., Belusic, D., Böing, S.J., Brooke, J.K., Harlow, R.C., Harris, P.P., Hunt, K.M., Jayakumar, A. and Mitra, A.K. (2019) A case-study of land–atmosphere coupling during monsoon onset in northern India. *Quarterly Journal of the Royal Meteorological Society*, 146, 2891–2905.

Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C., Edwards, J., Hendry, M., Porson, A. and Gedney, N. (2011) The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. *Geoscientific Model Development*, 4, 677–699.

Bhat, G.S., Morrison, R., Taylor, C.M., Bhattacharya, B.K., Paleri, S., Desai, D., Evans, J.G., Pattnaik, S., Sekhar, M., Nigam, R., Sattar, A., Angadi, S.S., Kacha, D., Patidar, A., Tripathi, S.N., Krishnan, K.V.N. and Sisodiya, A. (2019) Spatial and temporal variability in energy and water vapour fluxes observed at seven sites on the Indian subcontinent during 2017. *Quarterly Journal of the Royal Meteorological Society*, 146, 2853–2866.

Birch, C.E., Parker, D.J., Marsham, J.H., Copsey, D. and Garcia-Carreras, L. (2014) A seamless assessment of the role of convection in the water cycle of the West African Monsoon. *Journal of Geophysical Research: Atmospheres*, 119, 2890–2912.

Birch, C.E., Roberts, M.J., Garcia-Carreras, L., Ackerley, D., Reeder, M.J., Lock, A.P. and Schiemann, R. (2015) Sea-breeze dynamics and convection initiation: the influence of convective parameterization in weather and climate model biases. *Journal of Climate*, 28, 8093–8108.

Boutle, I.A., Abel, S., Hill, P. and Morcrette, C.J. (2014) Spatial variability of liquid cloud and rain: observations and microphysical effects. *Quarterly Journal of the Royal Meteorological Society*, 140, 583–594.

Brown, A.R., Milton, S.B., Cullen, M.J.P., Golding, B., Mitchell, J. and Shelly, A. (2012) Unified modeling and prediction of weather and climate: a 25-year journey. *Bulletin of the American Meteorological Society*, 93, 1865–1877.

Chakraborty, A. and Agrawal, S. (2017) Role of West Asian surface pressure in summer monsoon onset over central India. *Environmental Research Letters*, 12. https://doi.org/10.1088/1748-9326/aa76ca

Chakraborty, A., Nanjundiah, R. and Srinivasan, J. (2006) Theoretical aspects of the onset of Indian summer monsoon from perturbed orography simulations in a GCM. *Annales Geophysicae*, 24, 2075–2089. https://doi.org/10.5194/angeo-24-2075-2006

Clark, D.B., Mercado, L.M., Sitch, S., Jones, C.D., Gedney, N., Best, M.J., Pryor, M., Rooney, G.G., Essery, R.L.H., Blyth, E., Boucher, O., Harding, R.J., Huntingford, C. and Cox, O.M. (2011) The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4, 701–722.

Cullen, M.J.P. (1993) The unified forecast/climate model. *Meteorological Magazine*, 122, 81–94.

Davies, T., Cullen, M.J.P., Malcolm, A.J., Mawson, M., Staniforth, A., White, A.A. and Wood, N. (2005) A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131, 1759–1782.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van der Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158.

Fletcher, J.K., Parker, D.J., Turner, A.G., Menon, A., Martin, G.M., Birch, C.E., Mitra, A.K., Mrudula, G., Hunt, K.M. and Taylor, C.M. (2020) The dynamic and thermodynamic structure of the monsoon over southern India: new observations from the INCOMPASS IOP. *Quarterly Journal of the Royal Meteorological Society*, 146, 2867–2890.

Gadgil, S. and Sajani, S. (1998) Monsoon precipitation in the AMIP runs. *Climate Dynamics*, 14, 659–689.

Gallego-Elvira, B., Taylor, C.M., Harris, P.P. and Ghent, D. (2019) Evaluation of regional-scale soil moisture–surface flux dynamics in Earth system models based on satellite observations of land surface temperature. *Geophysical Research Letters*, 46, 5480–5488.

Gentine, P., Entekhabi, D. and Polcher, J. (2011) The diurnal behavior of evaporative fraction in the soil–vegetation–atmospheric boundary-layer continuum. *Journal of Hydrometeorology*, 12, 1530–1546.

Hohenegger, C. and Stevens, B. (2013) Preconditioning deep convection with cumulus congestus. *Journal of the Atmospheric Sciences*, 70, 448–464.

Holloway, C., Woolnough, S. and Lister, G. (2012) Precipitation distributions for explicit versus parametrized convection in a large-domain high-resolution tropical case-study. *Quarterly Journal of the Royal Meteorological Society*, 138, 1692–1708.

Hunt, K.M. and Turner, A.G. (2017) The effect of soil moisture perturbations on Indian monsoon depressions in a numerical weather prediction model. *Journal of Climate*, 30, 8811–8823.

IMD (2016) *2016 Southwest Monsoon End of Season Report*. New Delhi: India Meteorological Department.

Ji, Y. and Vernekar, A.D. (1997) Simulation of the Asian summer monsoons of 1987 and 1988 with a regional model nested in a global GCM. *Journal of Climate*, 10, 1965–1979.

Johnson, R.H., Ciesielski, P.E. and Hart, K.A. (1996) Tropical inversions near the 0 °C level. *Journal of the Atmospheric Sciences*, 53, 1838–1855.

Johnson, S.J., Levine, R.C., Turner, A.G., Martin, G.M., Woolnough, S.J., Schiemann, R., Mizielinski, M.S., Roberts, M.J., Vidale, P.L., Demory, M.-E. and Strachan, J. (2016) The resolution sensitivity of the South Asian monsoon and Indo-Pacific in a global 0.35° AGCM. *Climate Dynamics*, 46, 807–831.

Joyce, R.J., Janowiak, J.E., Arkin, P.A. and Xie, P. (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5, 487–503.

Karmacharya, J., Levine, R., Jones, R., Moufouma-Okia, W. and New, M. (2015) Sensitivity of systematic biases in South Asian summer monsoon simulations to regional climate model domain size and implications for downscaled regional process studies. *Climate Dynamics*, 45, 213–231.

Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E. and Lawrence, D. (2004) Regions of strong coupling between soil moisture and precipitation. *Science*, 305, 1138–1140.

Krishnamurti, T., Simon, A., Thomas, A., Mishra, A., Sikka, D., Niyogi, D., Chakraborty, A. and Li, L. (2012) Modeling of forecast sensitivity on the march of monsoon isochrones from Kerala to New Delhi: the first 25 days. *Journal of the Atmospheric Sciences*, 69, 2465–2487.

Kumar, K.K., Kumar, K.R., Ashrit, R., Deshpande, N. and Hansen, J. (2004) Climate impacts on Indian agriculture. *International Journal of Climatology*, 24, 1375–1393.

Kumar, K.R., Sahai, A., Kumar, K.K., Patwardhan, S., Mishra, P., Revadekar, J., Kamala, K. and Pant, G. (2006) High-resolution climate change scenarios for India for the 21st century. *Current Science*, 90, 334–345.

Levine, R.C. and Martin, G.M. (2018) On the climate model simulation of Indian monsoon low pressure systems and the effect of remote disturbances and systematic biases. *Climate Dynamics*, 50, 4721–4743.

Levine, R.C. and Turner, A.G. (2012) Dependence of Indian monsoon rainfall on moisture fluxes across the Arabian Sea and the impact of coupled model sea surface temperature biases. *Climate Dynamics*, 38, 2167–2190.

Levine, R.C., Turner, A.G., Marathayil, D. and Martin, G.M. (2013) The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall. *Climate Dynamics*, 41, 155–172.

Lock, A.P., Brown, A.R., Bush, M.R., Martin, G.M. and Smith, R.N.B. (2000) A new boundary-layer mixing scheme Part I: scheme description and single-column model tests. *Monthly Weather Review*, 128, 3187–3199.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W. (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21, 1303–1330.

Madden, R.A. and Robitaille, F. (1970) A comparison of the equivalent potential temperature and the static energy. *Journal of the Atmospheric Sciences*, 27, 327–329.

Marsham, J.H., Dixon, N.S., Garcia-Carreras, L., Lister, G.M., Parker, D.J., Knippertz, P. and Birch, C.E. (2013a) The role of moist convection in the West African monsoon system: insights from continental-scale convection-permitting simulations. *Geophysical Research Letters*, 40, 1843–1849.

Martin, G.M., Milton, S.B., Senior, C.A., Brooks, M., Ineson, S., Reichler, T. and Kim, J. (2010) Analysis and reduction of

systematic errors through a seamless approach to modeling weather and climate. *Journal of Climate*, 23, 5933–5957.

Martin, G.M., Klingaman, N.P. and Moise, A.F. (2017) Connecting spatial and temporal scales of tropical precipitation in observations and the MetUM-GA6. *Geoscientific Model Development*, 10, 105–126.

Martínez-de la Torre, A., Blyth, E.M. and Robinson, E.L. (2019) Evaluation of drydown processes in global land surface and hydrological models using flux tower evapotranspiration. *Water*, 11, 356.

Menon, A., Turner, A., Martin, G.M. and MacLachlan, C. (2018) Modelling the moistening of the free troposphere during the northwestward progression of Indian monsoon onset. *Quarterly Journal of the Royal Meteorological Society*, 144, 1152–1168.

Mitra, A.K., Bohra, A., Rajeevan, M. and Krishnamurti, T. (2009) Daily Indian precipitation analysis formed from a merge of rain-gauge data with the TRMM TMPA satellite-derived rainfall estimates. *Journal of the Meteorological Society of Japan. Series II*, 87, 265–279.

Pai, D., Bhate, J., Sreejith, O. and Hatwar, H. (2011) Impact of MJO on the intraseasonal variation of summer monsoon rainfall over India. *Climate Dynamics*, 36, 41–55.

Parker, D.J., Willetts, P., Birch, C.E., Turner, A.G., Marsham, J.H., Taylor, C.M., Kolusu, S. and Martin, G.M. (2016) The interaction of moist convection and mid-level dry air in the advance of the onset of the Indian monsoon. *Quarterly Journal of the Royal Meteorological Society*, 142, 2256–2272.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C. and Defourny, P. (2015) Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative. *Geoscientific Model Development*, 8, 2315–2328.

Prasad, O., Singh, O. and Prasad, K. (2018) 2016 Southwest monsoon and its long-range forecast. *Journal of Indian Geophysical Union*, 22, 90–100.

Prein, A.F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O. and Feser, F. (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Reviews of Geophysics*, 53, 323–361.

Randall, D., Khairoutdinov, M., Arakawa, A. and Grabowski, W. (2003) Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society*, 84, 1547–1564.

Rieck, M., Hohenegger, C. and van Heerwaarden, C.C. (2014) The influence of land surface heterogeneities on cloud size development. *Monthly Weather Review*, 142, 3830–3846.

Shao, X., Li, S., Liu, N. and Song, J. (2018) The Madden–Julian Oscillation during the 2016 summer and its possible impact on rainfall in China. *International Journal of Climatology*, 38, 2575–2589.

Smith, R.N.B. (1990) A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, 116, 435–460.

Sperber, K. and Palmer, T.N. (1996) Interannual tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project. *Journal of Climate*, 9, 2727–2750.

Sperber, K.R., Annamalai, H., Kang, I.-S., Kitoh, A., Moise, A., Turner, A., Wang, B. and Zhou, T. (2013) The Asian summer monsoon: an intercomparison of CMIP5 versus CMIP3 simulations of the late 20th century. *Climate Dynamics*, 41, 2711–2744.

Stirling, A. and Stratton, R.A. (2012) Entrainment processes in the diurnal cycle of deep convection over land. *Quarterly Journal of the Royal Meteorological Society*, 138, 1135–1149.

Stratton, R.A., Senior, C.A., Vosper, S.B., Folwell, S.S., Boutle, I.A., Earnshaw, P.D., Kendon, E., Lock, A.P., Malcolm, A. and Manners, J. (2018) A pan-African convection-permitting regional climate simulation with the Met Office Unified Model: CP4-Africa. *Journal of Climate*, 31, 3485–3508.

Taylor, C.M., de Jeu, R.A., Guichard, F., Harris, P.P. and Dorigo, W.A. (2012) Afternoon rain more likely over drier soils. *Nature*, 489, 423.

Taylor, C.M., Birch, C.E., Parker, D.J., Dixon, N., Guichard, F., Nikulin, G. and Lister, G.M. (2013) Modeling soil moisture–precipitation feedback in the Sahel: importance of spatial scale versus convective parameterization. *Geophysical Research Letters*, 40, 6213–6218.

Turner, A.G., Bhat, G., Martin, G.M., Parker, D.J., Taylor, C., Mitra, A.K., Tripathi, S.N., Milton, S.B., Rajagopal, E. and Evans, J.G. (2020) Interaction of convective organization with monsoon precipitation, atmosphere, surface and sea: the 2016 INCOMPASS field campaign in India. *Quarterly Journal of the Royal Meteorological Society*, 146, 2828–2852.

Van Genuchten, M.T. (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44, 892–898.

Volonté, A., Turner, A.G. and Menon, A. (2020) Airmass analysis of the processes driving the progression of the Indian summer monsoon. *Quarterly Journal of the Royal Meteorological Society*, 146, 2949–2980.

Wang, B. (2005). Theory, pp. 315–317 In: *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, editor=Lau, K.M., editor=Waliser, D.E. Berlin: Springer.

Wheeler, M.C. and Hendon, H.H. (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Monthly Weather Review*, 132, 1917–1932.

Willetts, P., Marsham, J.H., Birch, C.E., Parker, D.J., Webster, S. and Petch, J. (2017) Moist convection and its upscale effects in simulations of the Indian monsoon with explicit and parametrized convection. *Quarterly Journal of the Royal Meteorological Society*, 143, 1073–1085.

Williams, K., Copsey, D., Blockley, E., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. and Hill, R. (2018) The Met Office global coupled model 3.0 and 3.1 (GC3. 0 and GC3. 1) configurations. *Journal of Advances in Modeling Earth Systems*, 10, 357–380.

Wilson, D.R., Bushell, A.C., Kerr-Munslow, A.M., Price, J.D. and Morcrette, C.J. (2008) Pc2: a prognostic cloud fraction and condensation scheme. I: scheme description. *Quarterly Journal of the Royal Meteorological Society*, 134, 2093–2107.

Wood, N., Staniforth, A., White, A.A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S.B. and Zerroukat, M. (2014) An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, 140, 1505–1520.

Woodhams, B.J., Birch, C.E., Marsham, J.H., Bain, C.L., Roberts, N.M. and Boyd, D.F. (2018) What is the added value of a convection-permitting model for forecasting extreme rainfall over tropical East Africa?. *Monthly Weather Review*, 146, 2757–2780.

Wu, R., Kirtman, B.P. and Pegion, K. (2006) Local air–sea relationship in observations and model simulations. *Journal of Climate*, 19, 4914–4932.

Yang, G.-Y. and Slingo, J.M. (2001) The diurnal cycle in the tropics. *Monthly Weather Review*, 129, 784–801.

How to cite this article: Menon, A., Turner, A.G., Volonté, A., Taylor, C.M., Webster, S. & Martin, G. (2021) The role of mid-tropospheric moistening and land-surface wetting in the progression of the 2016 Indian monsoon. *Quarterly Journal of the Royal Meteorological Society*, 1–23. Available from: https://doi.org/10.1002/qj.4183